PROCEEDINGS OF
SYMPOSIA IN APPLIED MATHEMATICS

Volume V

5C

PROCEEDINGS OF
SYMPOSIA IN APPLIED MATHEMATICS
VOLUME V

WAVE MOTION AND VIBRATION THEORY

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK TORONTO LONDON
1954

FOR THE AMERICAN MATHEMATICAL SOCIETY
80 WATERMAN STREET, PROVIDENCE, RHODE ISLAND

PROCEEDINGS OF THE
FIFTH SYMPOSIUM IN APPLIED MATHEMATICS
OF THE AMERICAN MATHEMATICAL SOCIETY

Held at Carnegie Institute of Technology
June 16–17, 1952

Albert E. Heins
EDITOR

EDITORIAL COMMITTEE

R. V. Churchill
Eric Reissner
A. H. Taub

Library of Congress Catalog Card Number: 50-1183

CONTENTS

EDITOR'S PREFACE

This volume contains the papers which were presented at the Fifth Symposium in Applied Mathematics of the American Mathematical Society held at the Carnegie Institute of Technology on June 16 and 17, 1952. The subject of the Symposium was *Wave Motion and Vibration Theory*, and the four sessions were devoted to *Stability of Fluid Motions, Hydrodynamic Waves, Diffraction and Scattering Problems*, and *Vibration Theory*.

One of the papers appears as an abstract, because of prior arrangements for publication.

All who participated in the Symposium are indebted to the McGraw-Hill Book Company, Inc., which, beginning with the Proceedings of the Symposium on Elasticity, has undertaken in these uncertain times the task of bringing the Proceedings of these Symposia on Applied Mathematics to the scientific public in book form.

The Editor gratefully acknowledges the invaluable help afforded by the Committee on Arrangements, consisting of W. M. Whyburn, P. Chiarulli, P. Gustafson, G. H. Handelman, R. C. Meacham, L. E. Malvern, D. Moskovitz, and E. A. Whitman. The generous financial support provided by the Carnegie Institute of Technology greatly enhanced the success of the Symposium.

ALBERT E. HEINS
Editor

HYDRODYNAMIC STABILITY[1]

BY

C. C. LIN

I. Introduction

1. General formulation of the stability problem. The existence of two types of motions of a viscous fluid—laminar and turbulent—immediately raises the question: Which type of motion is the more likely to occur? It has now been generally recognized that turbulent motion is the more natural state of fluid motion, and laminar motion occurs only when the Reynolds number is so low that deviations from it are liable to be damped out. For certain types of flows, it has been found possible to keep the flow laminar for higher and higher Reynolds numbers by keeping the disturbance smaller and smaller. The question then may be asked, for a given flow: Is it stable relative to infinitesimally small disturbances? This is the problem of hydrodynamic stability.

Mathematically, the problem is as follows: Suppose the system of hydrodynamic equations has a time-independent solution

$$\bar{u}_i(x_k),\ \bar{p}(x_k),\ \bar{T}(x_k) \tag{1}$$

for the components of velocity, pressure, and temperature. Consider an initial-value problem with these variables slightly different from this steady-state solution. If the solution approaches solution (1) as time $t \to \infty$, the motion is stable. Otherwise, it is unstable. It should be noted that instability does not necessarily lead to turbulent motion; it could lead to another state of laminar motion.

To solve the problem of hydrodynamic stability, one must therefore follow the solution of a system of nonlinear (quasi-linear) partial differential equations, and the task is in general very difficult. In the usual approaches to the problem, the mathematical formulation is cast in a different way. We *assume* that, for small disturbances, the equations for the small disturbances may be *linearized; i.e.*, terms quadratic or higher in the disturbances and their derivatives may be rejected. The resultant linear system of equations contains time t only through derivatives with respect to t, and hence solutions containing an exponential time factor $e^{\sigma t}$ may be expected. We then have a characteristic-value problem, with σ as the parameter. If all the characteristic values of σ have negative real parts, the motion is stable with respect to infinitesimal disturbances. If some of the characteristic values σ have positive real parts, the motion is said to be unstable. The argument for this statement is briefly

[1] Some material was inserted after the presentation of this paper at the meeting. This added work was done with the assistance of Contract No. N5 ori-07872 of the Office of Naval Research.

as follows: If the flow is given a disturbance, finite or infinitesimal, we may conclude that it will, in general, not die out completely. For if a disturbance is to die out, it must first become so small that the linearized equations are applicable. The small disturbance may in general be expected to contain an unstable mode, and it therefore cannot disappear completely.

The validity of the process of linearization has sometimes been questioned in the problem of hydrodynamic stability, even though it is an often used process. In the literature, one finds claims of proof of complete stability in the nonlinear theory in cases where the linear theory definitely indicates instability. Careful examination of these "proofs" indicates that the claims are not justified.

Even if the linearization is accepted, very complicated mathematical problems often have to be clarified before the characteristic-value problem can be solved. It is therefore no great wonder that there has been a great deal of disagreement among workers in this field. Some of them, unfortunately, result from misunderstandings. In the present article, we shall try to present a discussion of some of these controversial points. The two following parts will outline the linearized theory and the conclusions derived therefrom. This will be followed by a discussion of the problem of friction layers and friction regions.

II. Stability of Flows Represented by Exact Solutions of Navier-Stokes Equations

2. Some exact solutions and their stability characteristics. To avoid the complication in the various approximations involved in the theory of the stability of the boundary layer, we shall first consider the stability problem for exact solutions. There are only a limited number of such solutions in the case of incompressible fluid, and the following types have been repeatedly studied for their stability characteristics:

(*a*) Poiseuille flow
(*b*) Couette flow
(*c*) Plane Poiseuille flow
(*d*) Plane Couette flow

No indication of instability has ever been found for cases (*a*) and (*d*). Case (*b*) is the classical case solved by G. I. Taylor [1], who calculated its stability characteristics and also confirmed his theoretical results by experiments. Some extension of the theoretical work in this case has been done by Meksyn [2]. It is recognized that in this case instability does not directly lead to turbulent motion.

The case of plane Poiseuille flow has been studied by many authors with contradictory conclusions. Heisenberg [3] was the first to reach the conclusion of the instability of the flow for sufficiently high Reynolds numbers, but he did not arrive at a critical value beyond which instability begins. The present writer carried out the detailed calculations [4], following the method

given by Heisenberg, and arrived at a Reynolds number of 5,300 based on the maximum velocity at the center of the channel and its half width (Fig. 1).

Recently, Pekeris [5] again concluded that the plane Poiseuille flow is stable with respect to small disturbances. The present writer believes that this conclusion arises from the fact that the formula he uses for the determination of characteristic values is essentially valid only in stable cases. It was decided to use an entirely different approach to settle this disagreement. Recently, L. H. Thomas [6] solved the characteristic-value problem by direct calculation with the help of the automatic computing machine. The results confirm those obtained from the Heisenberg method.

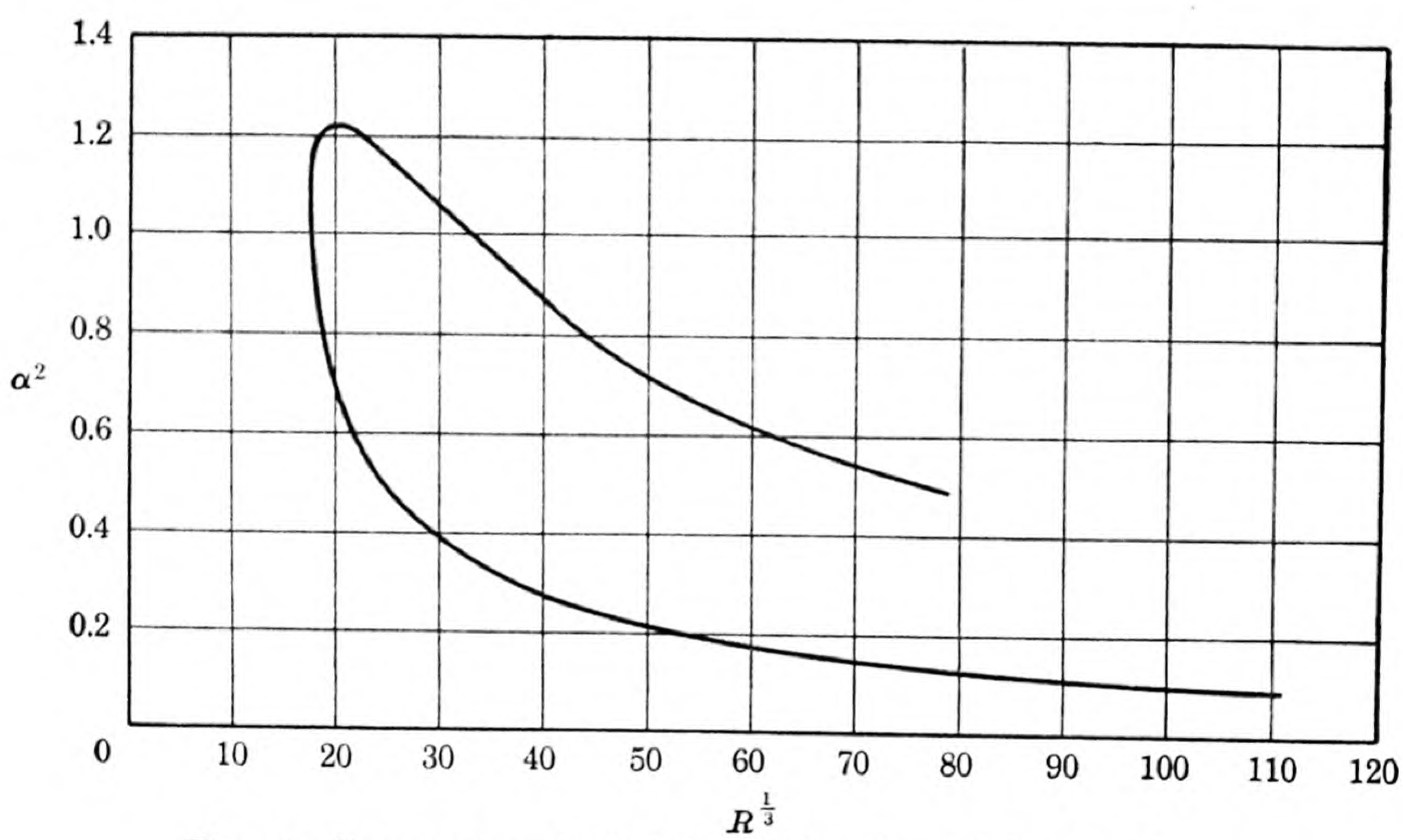


FIG. 1. Curve of neutral stability for plane Poiseuille flow.

3. Further discussion of the case of plane Poiseuille flow. Since the methods for dealing with this classical stability problem are representative of the methods used in the other cases considered below, we shall present the theory in some detail.

The basic flow has velocity components

$$\bar{u}_1 = \bar{u}_1(x_2), \qquad \bar{u}_2 = 0, \qquad \bar{u}_3 = 0. \tag{2}$$

It can be easily verified that the linearized equations for a three-dimensional disturbance contain variables x_1, x_3, and t only in the form of partial derivatives. Solutions containing the exponential factor $\exp i\{\alpha_1 x_1 + \alpha_3 x_3 + \sigma t\}$ may therefore be expected. In fact, if the disturbance is to be finite at infinity, it must be spatially *periodic* in x_1 and x_3.

Now, these disturbances represent waves propagating in the direction of $(\alpha_1,0,\alpha_3)$, which makes an angle

$$\tau = \arctan \frac{\alpha_3}{\alpha_1} \tag{3}$$

with the x_1-axis. If we introduce a rotation of coordinates in the x_1x_3-plane and take the new x_1^*-axis in this direction of wave propagation, the components of velocity of the basic flow would be

$$\bar{u}_1^* = \bar{u}_1(x_2) \cos \tau, \qquad \bar{u}_2^* = 0, \qquad \bar{u}_3^* = -\bar{u}_1(x_2) \sin \tau. \tag{4}$$

The disturbance is now propagating in the direction of the x_1^*-axis and independent of the variable x_3^*. In fact, it can easily be verified that, in this coordinate system, the equations governing the disturbances $(u_1^{*\prime}, u_2^{*\prime}, p')$ would be independent of $\bar{u}_3^*$ and $u_3^{*\prime}$. Thus, we have essentially a two-dimensional disturbance corresponding to the basic velocity components

$$\bar{u}_1^* = \bar{u}_1(x_2) \cos \tau, \qquad \bar{u}_2^* = 0, \qquad \bar{u}_3^* = 0. \tag{5}$$

It is thus the same as the basic flow in the original problem, except that the velocity, and hence the Reynolds number, is reduced by the factor $\cos \tau$. Thus, it is only necessary to consider two-dimensional disturbances in the linearized theory, especially in the determination of the minimum critical Reynolds number. A somewhat different way of deriving this result was first given by Squire [7].

Notice that the conclusion was based on the linearized equations. It is well known that three-dimensional motions are important in fully developed turbulence and hence may be expected to be important when the nonlinear terms are considered.

The linearized differential equation for the amplitude function of two-dimensional disturbances is (in dimensionless form)

$$\phi^{\mathrm{iv}} - 2\alpha^2\phi'' + \alpha^4\phi + i\alpha R[(w - c)(\phi'' - \alpha^2\phi) - w''\phi] = 0, \tag{6}$$

where the disturbance is given by the stream function

$$\psi = \phi(y) \exp [i\alpha(x - ct)],$$

$w(y)$ is the basic velocity distribution in dimensionless form,

$$w(y) = 1 - y^2,$$

and R is the Reynolds number of the basic flow. The boundary planes of the flow are given by $y = \pm 1$, where the conditions

$$\phi(\pm 1) = 0, \qquad \phi'(\pm 1) = 0 \tag{7}$$

must be satisfied. We have thus a characteristic-value problem resulting in a condition of the type

$$F(\alpha, R, c) = 0. \tag{8}$$

For each pair of real values α and R, there is a characteristic value c, in general complex. If the imaginary part c_i of c is positive, the disturbance is unstable according to the linearized theory. If $c_i < 0$, the disturbance is damped. If $c_i = 0$, the disturbance is a sustained oscillation.

The characteristic-value problem can be solved by numerical integration. However, because instability may be expected only for large values of the Reynolds number, the solution varies very rapidly in y and very fine steps must be taken. The control of errors also becomes a difficult problem. These difficulties have, however, been overcome by L. H. Thomas. It is estimated by him that the amount of work carried out would require 100 years by hand computation. It took 2 weeks on the high-speed IBM electronic calculator. As mentioned before, the results agree remarkably well with those obtained from the method of asymptotic solution of (6), which we shall describe presently. The method of numerical calculation is at present limited to relatively low Reynolds numbers almost barely sufficient to produce instability for a limited range of values of α. Advance knowledge of the analytical results helps to give the correct combinations of (α,R) and to ensure that the calculations would not be fruitless.

To solve the characteristic-value problem, Heisenberg obtained four solutions by asymptotic methods. Two of these are given by writing

$$\phi = \phi^{(0)}(y) + \frac{1}{\alpha R}\phi^{(1)}(y) + \cdots, \tag{9}$$

where $\phi^{(0)}$ is the solution of the inviscid equation

$$(w - c)(\phi'' - \alpha^2\phi) - w''\phi = 0. \tag{10}$$

Let us denote the solutions by $\phi_1(y)$ and $\phi_2(y)$. The other two are given by the formulas of the type

$$\phi = \left(f_0(y) + \frac{1}{\sqrt{\alpha R}} f_1(y) + \cdots\right) \exp\left[\pm \int \sqrt{i\alpha R(w - c)}\, dy\right], \tag{11}$$

which may be denoted by $\phi_3(y)$ and $\phi_4(y)$, where ϕ_3 shall be taken to be the solution which decreases exponentially as we move away from the solid boundary. The characteristic-value problem is also split into two parts, by using the symmetry property of $w(y)$:

$$\begin{array}{lllll} (12) & (a)\ \phi(-1) = 0, & \phi'(-1) = 0, & \phi(0) = 0, & \phi''(0) = 0; \\ & (b)\ \phi(-1) = 0, & \phi'(-1) = 0, & \phi'(0) = 0, & \phi'''(0) = 0. \end{array}$$

Instability is found only in the latter case. In fact, if Φ is the solution of the inviscid equation satisfying the condition

$$\Phi'(0) = 0, \tag{13}$$

then the equation defining the characteristic values is simply

$$\frac{\Phi'(-1)}{\Phi(-1)} = \frac{\phi_3'(-1)}{\phi_3(-1)}. \tag{14}$$

However, form (11) for ϕ_3 is not adequate for yielding all the results required. Another form of the solution, expressed in terms of Hankel functions, is

required. This is obtained by introducing the new variable

$$\eta = \frac{y - y_c}{\epsilon}, \qquad \epsilon = (\alpha R)^{-} \tag{15}$$

and expressing the solution in the form

$$\phi = \chi^{(0)}(\eta) + \epsilon\chi^{(1)}(\eta) + \cdots . \tag{16}$$

The procedure is suggested by considering the order of magnitude of the two terms ϕ^{iv} and $i\alpha R(w - c)\phi''$ in the neighborhood of the point y_c, where $w = c$. This method actually yields four solutions at once, and the particular solution ϕ_3 is given by

$$\phi_3 = \int_{+\infty}^{\zeta} d\zeta \int_{+\infty}^{\zeta} H_{\frac{1}{3}}^{(1)}\left[\frac{3}{2}(i\zeta)^{\frac{3}{2}}\right]\zeta^{\frac{1}{2}}\, d\zeta; \qquad \zeta = [w'(y_c)]^{\frac{1}{3}}\eta. \tag{17}$$

Heisenberg did not carry out the calculations with this new form, but he did suggest its use and indicated the kind of result that may be expected, including the correct general shape of the curve of neutral stability. The detailed calculations were carried out by the present author.

4. Mathematical problems associated with the asymptotic solution of the Orr-Sommerfeld equation. The above analysis is not free from dubious points, which will be discussed presently. Strictly speaking, such points can be settled only after a thorough investigation of the way in which the asymptotic solutions represent approximations to the exact solutions of (6). Such a complete theory was developed by Wasow [8]. Earlier, Tollmien [9] gave a satisfactory solution in the case of real variables. Other workers were also able to arrive at correct answers without the benefit of the complete mathematical theory.

(*a*) The first question that may be raised is the following: While equation (6) should have single-valued solutions, the asymptotic solutions (9) and (11) are multiple-valued, with a branch point at $y = y_c$ (which will be referred to as the critical point). The question arises: Which branches of these multiple-valued solutions are correct? A practical form of the answer is as follows: If the value of $w'(y)$ is essentially *positive* at the critical point, the path for determining the correct branch of these solutions should be taken *below* the critical point; if *negative, above* it. The rigorous form of the answer is expressed in terms of the real part of the integral

$$Q = \int_{y_c}^{y} \sqrt{i(w - c)}\, dy. \tag{18}$$

If one follows a path along which $Rl(Q)$ is monotonically increasing, the asymptotic solutions (9) and (11) would remain valid along the whole path (see Fig. 2).

(*b*) The second question is the following: Solutions of the type (9) and (11) are adequate for fixed values of y while αR becomes large. On the other hand,

the type of solution given by (16) is adequate for fixed η, as ϵ becomes small (αR becomes large). Can these two types of solutions be used at the same time?

Formally, one can get some assurance that this is correct by expanding solutions (16) asymptotically and comparing them with solutions (9) and (11). But this is not strictly satisfactory from the mathematical point of view. In particular, the asymptotic expansion of (17) does not agree with the asymptotic solution (11) except in the immediate neighborhood of y_c. This can actually be remedied [9] by redefining ζ in (17) by

$$\zeta = \left[\frac{3}{2}\int_{y_c}^{y} \sqrt{\alpha R(w - c)}\, dy\right]^{\frac{2}{3}} \tag{19}$$

and attaching a suitable factor to (17). Numerically, this modification turns out to be not significant in the present case. We may, however, note that this type of modification will be important for supersonic boundary layers.

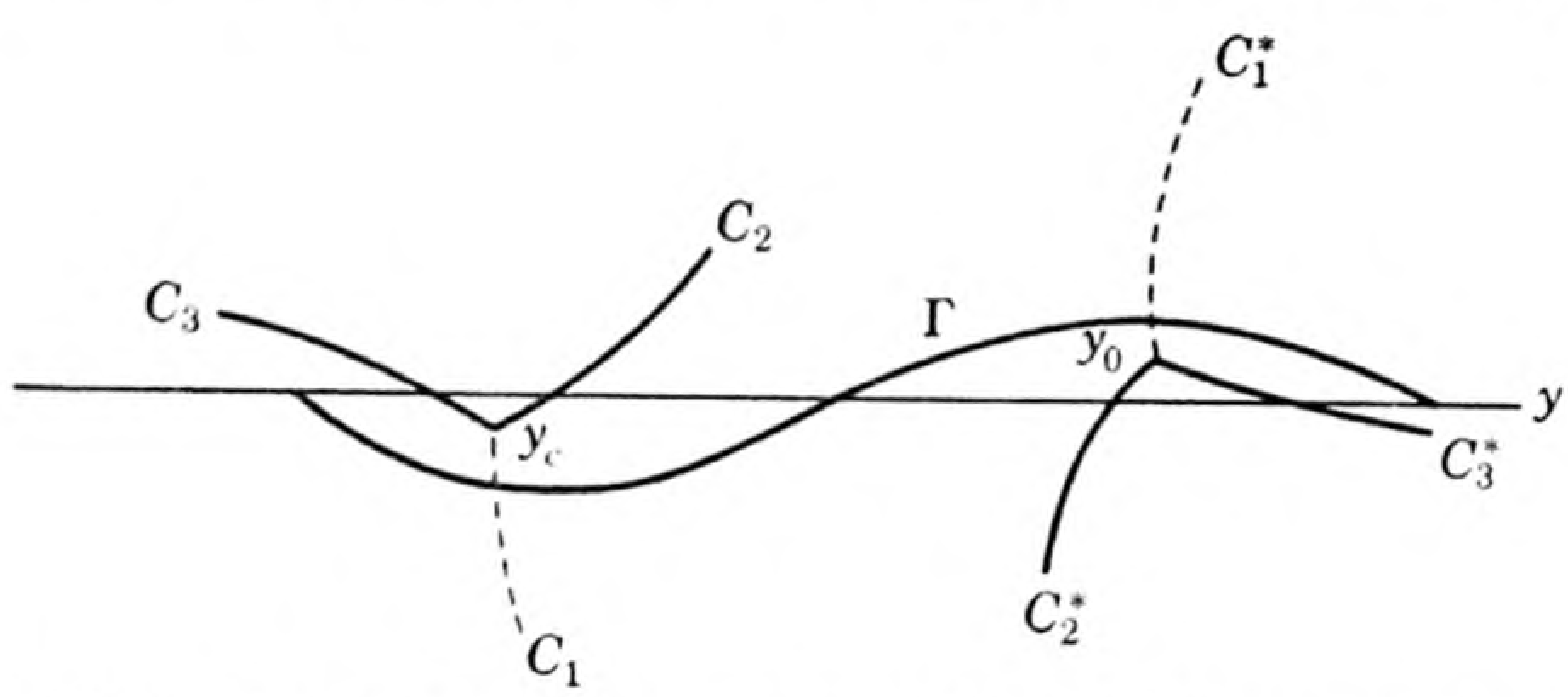


FIG. 2. Typical path Γ along which the asymptotic solutions remain valid. $Rl(Q) = 0$ along C_1, C_2, C_3; $Rl(Q) =$ constant along C_1^*, C_2^*, C_3^*.

The real difficulty arises with one of the solutions of equation (10) (often denoted by ϕ_2). The application of the usual theory of differential equations to the study of the nature of the solution of (10) in the neighborhood of $y = y_c$ was made by Tollmien [10]. It is found that one of the solutions is regular, the other has a logarithmic singularity with derivative becoming infinite as $y \to y_c$. These solutions were given by Tollmien in the form

$$\begin{aligned} \phi_1 &= y + ay^2 + \cdots, \\ \phi_2 &= 1 + by^2 + \cdots + \frac{w_c''}{w_c'}\phi_1 \log y, \end{aligned} \tag{20}$$

where a and b are constants and $w_c' = w'(y_c)$, $w_c'' = w''(y_c)$. The inviscid solution ϕ_2 obviously fails to represent a solution of (6) in the immediate neighborhood of $y = y_c$, and one may expect viscous effects to be called into play. The point $y = y_c$ is often referred to as the "inner-friction layer." This would have made no difference to the boundary-value problem, if the boundary points were all at finite distances (on the y-scale) from the point y_c. Actually, however, the calculations of the neutral curve in the present case indicate that this distance approaches zero on the y-scale if $\alpha R \to \infty$. In fact,

for the lower branch of the neutral curve, it is always at a finite distance on the η-scale. This means that there is no guarantee that asymptotic solutions of the type (9) are valid for use with the boundary-value problem. (Indeed, the asymptotic solutions (11) *are* not valid for the computation of the lower branch of the neutral curve. For this reason Heisenberg suggested the use of (17).) However, an estimate of the error in ϕ_2 and ϕ_2' (which are needed in (12)[2]) can be obtained from the recent work of Tollmien [9] (for real y and real parameters α, R, and c), and Wasow [8] (for general values of y and the parameters). It can be shown that for $y - y_c = 0(\epsilon)$, that is, finite η, the percentage error in ϕ_2 is of the order of ϵ and that in ϕ_2' is of the order of $1/\log \epsilon$ with a small coefficient. This explains why the use of (10) gives the correct approximate answer. Tollmien also gave an improved solution for real c with percentage error of the order of ϵ in ϕ_2'. Within this error, Wasow's corresponding solution, for general values of c, is in agreement with Tollmien's in the special case. These solutions are, however, somewhat complicated to use for actual calculations.

(*c*) There is a third problem which is not of direct interest for the calculations of the neutral curve discussed above but which is of great interest in the general theory. What is the behavior of the solution in the limit $\alpha R \to \infty$? Does it go into a solution of the inviscid equation (10), satisfying inviscid boundary conditions $\phi(-1) = 0$, $\phi'(0) = 0$? This question will be discussed more fully in Part IV.

III. Stability of Flows Approximately Two-dimensional and Parallel

5. General remarks. In the present section, we shall chiefly be concerned with the stability of basic flows which are parallel or nearly parallel. The principal types of exactly parallel flows are those mentioned in Sec. 2. However, if one considers flows which are nearly parallel, there is the large class of flows of the boundary-layer type. Indeed, this generalization opens up a whole field of studies of the stability problems in a compressible fluid. Another striking example is the stability of the jet stream in the upper atmosphere. However, in any such discussion it is necessary to justify the application of the theory of stability of parallel flows. While it has been found that this theory is applicable to boundary-layer flow over a flat plate and over a convex surface, the stability characteristics of the boundary layer over a concave surface and that of the pressure flow through a slightly curved channel are chiefly controlled by disturbances of the type found by G. I. Taylor in Couette flow [1]. In either case, the basic flow is, however, not much influenced by the curvature of the solid boundaries.

As in the discussion in Sec. 3, the equations for small periodic three-dimensional disturbances can be reduced to those for two-dimensional disturbances. So can the boundary conditions. The transformation in the case of the

[2] The point $y = 0$ is far away from the point $y = y_c$, and hence derivatives needed at that point are not subjected to the above complication.

boundary layer in a compressible fluid has been given recently by Dunn and the present writer [11]. However, there are two difficulties which prevent the neglect of the effect of three-dimensional disturbances, as is done in the case of plane Poiseuille flow. First, in the general case, the velocity distribution could change from station to station as we move downstream. Thus, although the transformation is valid from point to point, the history of a three-dimensional disturbance has no obvious correspondence with that of a two-dimensional disturbance. Second, in the case of the boundary layer in a compressible fluid, the temperature and velocity distribution in the equivalent two-dimensional problem does not correspond to that of an ordinary boundary layer. Thus, while the solution of the mathematical problem for a three-dimensional disturbance can be treated by the same method as for a two-dimensional one, there is no simple physical interpretation of the correspondence property.

With the introduction of general profiles, the Rayleigh-Tollmien criterion for the stability of parallel flows takes on added significance. This criterion is as follows: If the velocity distribution does not show a point of inflection, the motion is stable without the influence of viscosity. On the other hand, under certain general restrictions, the existence of the point of inflection implies instability. Such a case does not exist if only the exact solutions discussed in Sec. 2 are considered.

The present writer interpreted the occurrence of the point of inflection as the extremum of vorticity [4] and described a physical process to interpret the Rayleigh-Tollmien criterion. Further discussions of this point will not be given in this paper.

Lees and the present writer [12] generalized the condition to the case of the compressible fluid. In this case, the condition is that the gradient of the product of density times vorticity shall vanish. No simple physical interpretation has yet been found.

There is one further complication in the case of nearly parallel flows. The relative importance of the departure from parallel flow as compared with the viscous effect is often something to be borne in mind. For example, in the case of pressure flow through a channel, the calculations of Dean [13] indicate that a slight curvature in the channel (with a radius of curvature of 100 channel widths, say) would cause the three-dimensional disturbance to occur at a lower Reynolds number than that associated with the stability of the straight-channel flow. In the case of flows of the boundary-layer type, the rate of growth of the boundary layer, and consequently the x-dependence of the main flow, is of the order of $1/R$, where R is the Reynolds number based on the boundary-layer thickness, occurring in the stability equation. If such a dependence is neglected, as is usually the case, there may be some doubt of the legitimacy of including the viscous terms in equation (9).

In the next three sections, we discuss three types of problems each with some distinctive characteristics. This is not intended to be a treatment of all the types of problems of consequence, as the previous discussions clearly indicate.

6. Boundary layer over a flat plate. The method of solution given originally by Heisenberg and described in Sec. 3 can be applied to the calculation of the stability characteristics of the boundary layer, the essential difference being a modification of the boundary conditions. Such calculations were performed by Tollmien [10] and Schlichting [14], with some improvement of the mathematical theory. In carrying out the calculation, Tollmien and Schlichting used solutions of the inviscid equation (10) in power series of y. In his problem, Heisenberg suggested series solutions in powers of the parameter α^2. The result of Tollmien has been reproduced by the present writer by using the form of solution suggested by Heisenberg. The experimental work of Schubauer and Skramstad [15] confirms the theoretical calculations.

Studies of the boundary layer in a compressible fluid have been carried out by Lees and the present writer [12]. The problems are more complicated than in the incompressible case. However, it is found that the boundary conditions on the temperature fluctuations do not enter into the characteristic-value problem, and the final calculations can be carried out in a manner somewhat similar to the incompressible case. The main difference between the two cases is that the disturbances can travel with a wave speed supersonic relative to the free stream. These disturbances are assumed to be unimportant for stability considerations. Subsonic disturbances, on the other hand, do not always exist when the free stream is at a high enough supersonic Mach number. In fact, Lees [16] was able to show that when heat is withdrawn from the wall at a sufficient rate, the boundary layer could be completely stabilized with respect to all subsonic disturbances. This has some experimental confirmation and has caused great interest and much further work.[3]

There is, however, one modification necessary to arrive at more reliable numerical results in all these calculations. At high free-stream speeds, the density distribution cannot be adequately approximated in a simple manner, as is done in the usual theory. This point is important especially in the case of high values of the wave speed c, which is the case for subsonic disturbances in a supersonic stream. It is then necessary to modify the solutions of the type (16) along the lines discussed in Sec. 4. This has been carried out by D. W. Dunn at the suggestion of the present writer. We record here only the solution which corresponds to (17) in the incompressible case. This is the one which will enter into the final calculations of the characteristic-value problem. The system of notation used is explained by Lees and Lin [12].

$$
\begin{aligned}
&f_3 = (\alpha R)^{-\frac{1}{4}} F_3(y) g_3(\zeta), \\
(21)\qquad &\left\{\begin{aligned}
\phi_3 &= (\alpha R)^{-\frac{1}{12}} \Phi_3(y) \int_{+\infty}^{\xi} g_3(\zeta)\, d\zeta, \\
\frac{\pi_3}{\gamma M^2} &= (\alpha R)^{-\frac{5}{12}} P_3(y) \int_{+\infty}^{\xi} g_3(\zeta)\, d\zeta, \\
\theta_3 &= 0,
\end{aligned}\right.
\end{aligned}
$$

[3] Recent work by Dunn and the present writer indicates that there is a class of three-dimensional disturbances which cannot be completely stabilized by cooling.

where

$$
(22)\qquad \left\{
\begin{aligned}
F_3(y) &= \sqrt{\frac{\pi\rho}{3\mu_1}}\left[\frac{i(w-c)}{\nu_1}\right]^{-\frac{1}{4}}\left(\frac{3}{2}\int_{y_c}^{y}\sqrt{\frac{w-c}{\nu_1}}\,dy\right)^{\frac{1}{2}},\\
\Phi_3(y) &= -i\left(\frac{w-c}{\nu_1}\right)^{-\frac{1}{2}}\zeta_1^{\frac{1}{2}}F_3(y),\\
P_3(y) &= -\alpha^2\mu_1\zeta_1F_3(y),\\
g_3(y) &= \int_{+\infty}^{\xi}\zeta^{\frac{1}{2}}H_{\frac{1}{3}}^{(1)}\left[\frac{2}{3}(i\zeta)^{\frac{3}{2}}\right]d\zeta,\\
\zeta &= (\alpha R)^{\frac{1}{3}}\zeta_1,\\
\zeta_1 &= \left(\frac{3}{2}\int_{y_c}^{y}\sqrt{\frac{w-c}{\nu_1}}\,dy\right)^{\frac{2}{3}}.
\end{aligned}
\right.
$$

7. Mixing of two streams and related problems. Somewhat different methods have to be used for the treatment of problems of "free" boundary layers, such as the flow in a jet, a wake, and the mixing region between two parallel streams. The main point is that the boundary conditions are to be applied at large distances from the central part of the flow field. Thus, asymptotic solutions of the form (11) can be used, and the boundedness of the solution at large distances immediately shows that they must be rejected. Thus, viscous effects can only be found through the higher approximations of solution (9). This makes the nature of the stability characteristics differ considerably from the previous case. Studies in this direction have been made by Chiarulli [17], Lessen [18], Pai [19], and the present writer [20]. The present writer considered the mixing of two streams in the compressible case. Here the limitation to subsonic disturbances leads immediately to the following condition for stability: If the difference in the velocity of the two streams exceeds the sum of their velocities of sound, no subsonic two-dimensional disturbances can exist.

8. Meteorological problems. One interesting application of the stability theory is the study of oscillations in the atmosphere. There is a strong westerly current, strongest at the tropopause, whose streamlines exhibit a wavy pattern. If one considers an idealized model of a zonal current independent of altitude, a two-dimensional model over a spherical earth can be studied in much the same way as small oscillations about an ordinary two-dimensional steady flow. In fact, after a Mercator's projection, the inviscid equation for the amplitude of the oscillation is exactly the same as (10), except that w'' is replaced by the gradient of vorticity, which now includes the normal component of the rotation of the earth. Furthermore, since we are dealing with vorticity over a spherical earth, the vorticity gradient is more complicated than the second derivative of the velocity distribution.

As shown by Foote and Lin [21], the stability of the flow depends on whether this gradient vanishes or not. This confirms the interpretation of the Rayleigh-Tollmien criterion given by the present writer [4] for the usual stability of two-dimensional parallel flows. Rossby [22] adopts the mechanism described in this connection for the interpretation of the movement of polar air masses.

Detailed discussions of the oscillations in the westerlies and their meteorological implications are given by Kuo [23].

IV. Friction Layers and Friction Regions

9. The limiting behavior as the Reynolds number becomes infinite. We shall now examine in considerable detail the problem raised in Sec. 4 regarding the behavior of the solution of the problem of hydrodynamic stability in the limit of infinite Reynolds number. This question was emphasized by Heisenberg [3]. It is complicated because a formal limiting process sometimes gives the correct answer and sometimes does not.

(a) *Formal discussion of the limiting process.* To demonstrate our point, we shall now carry through a formal discussion of the limiting process and indicate where such discussions lead us into difficulties. For convenience, we shall refer to the equations in Sec. 3, but we have in mind other cases which can be handled with very little modification. For example, the case of the boundary layer requires only the replacement of condition (13) by the condition

$$\lim_{y\to\infty} \frac{\Phi'(y)}{\Phi(y)} = -\alpha.$$

If we neglect the effect of viscosity in (6), we arrive at equation (10), which should be solved with two boundary conditions. One of these is condition (13). The other is a condition at the solid wall $y = -1$. Since viscosity has been neglected, this condition is obviously

$$\phi(-1) = 0$$

corresponding to $u_2' = 0$. In a crude manner, this can be easily justified from (14), since

$$\lim_{\alpha R\to\infty} \frac{\phi_3(-1)}{\phi_3'(-1)} = 0,$$

whether ϕ_3 is given in form (11) or (17). However, if Φ' becomes infinite as $\alpha R \to \infty$, further investigation is needed.

Supposing that the inviscid formulation can be justified (as is the case for the neighborhood of the first branch of the neutral curve for the boundary layer under an adverse pressure gradient), we have a well-defined characteristic-value problem, except for the fact that there is a singularity of equation (10) at $w = c$ (if $w'' \neq 0$ at the same point). Thus, when viscous effects are considered, they are expected to play a role at the wall and possibly at the critical point $w = c$. From the form of solutions (11) and (17), it can be surmised that the outer friction layer at the wall has a thickness at most of the order of $(\alpha R)^{-\frac{1}{3}}$, which approaches zero as the Reynolds number becomes infinite. Such a layer is therefore quite similar to the usual boundary layer of Prandtl.

In the neutral case, the critical point $w = c$ lies on the real axis, generally corresponding to an interior layer of the fluid. Hence, we may expect to have an inner friction layer in this case. In the case of amplified and damped

disturbances, equation (10) has no singularity on the real axis. One might then argue that there is no inner friction layer in these cases. In fact, equation (10) indicates clearly that, along the real axis, amplified and damped disturbances are complex conjugates. This means that to each amplified solution there corresponds a damped solution for the same value of α. This is difficult to reconcile with the first branch of the neutral curve (Fig. 3) for the boundary layer under an adverse pressure gradient. There are amplified solutions for $\alpha < \alpha_s$ and damped solutions for $\alpha > \alpha_s$. By the complex-conjugate relation, there would be both amplified and damped disturbances for values of α larger and smaller than α_s. How can we then expect the neutral curve to separate the stable and unstable regions in the (α,R)-plane?

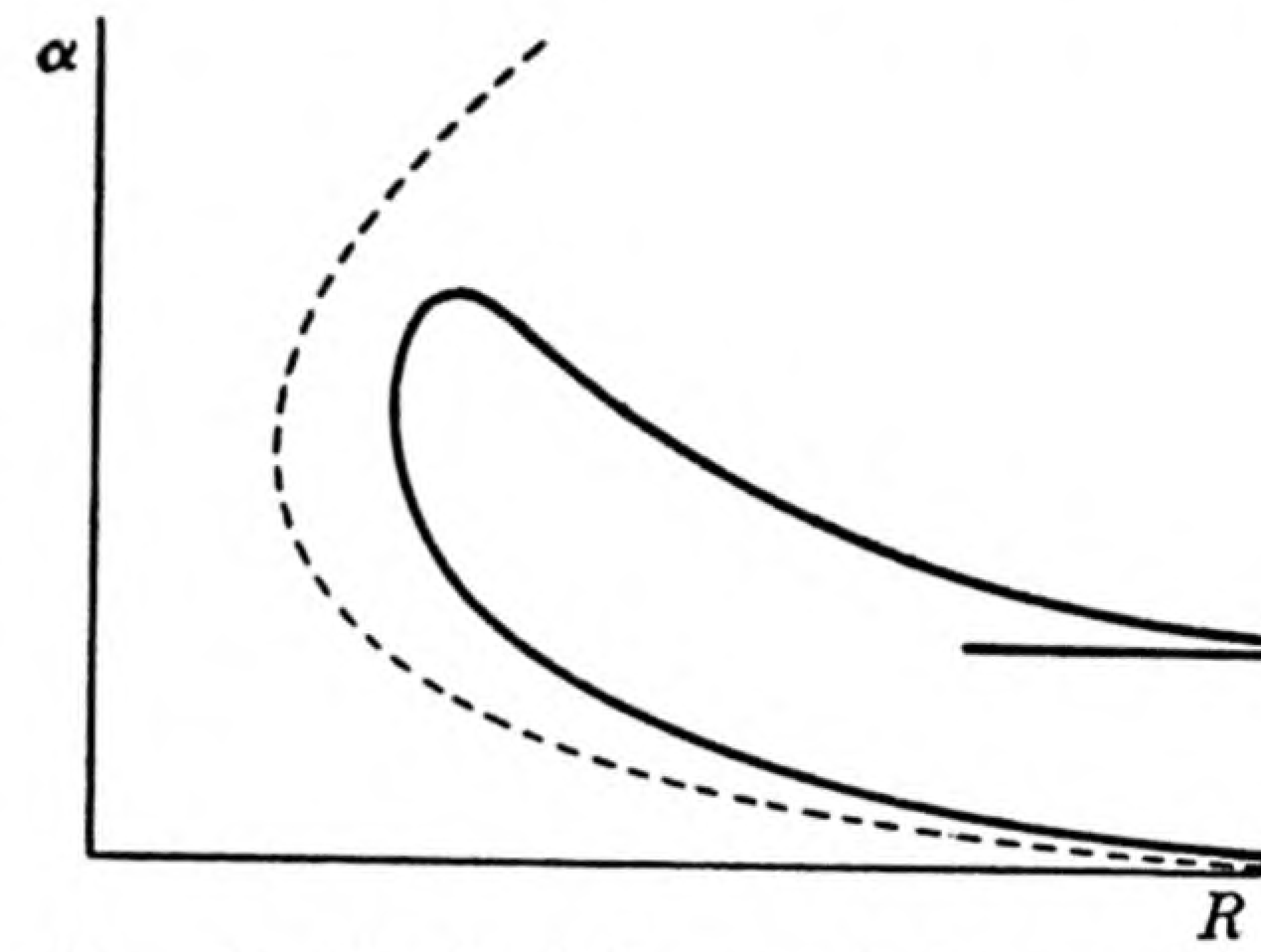


FIG. 3. Typical form of the neutral curve for boundary layer under an adverse pressure gradient; first branch has an asymptotic $\alpha_s \neq 0$.

(b) *Solution of the above dilemma.* This difficulty disappears if we accept the following conclusions, first reached by the present writer in 1944.

1. In the case of finite amplification, the solution $\phi(y)$ of the characteristic-value problem tends to a solution of the inviscid equation, throughout the part of the real axis corresponding to the field of flow, as the Reynolds number becomes infinite. Physically, this is found to mean that the effect of viscosity disappears completely in the interior of the fluid in the limit. There is no inner friction layer at sufficiently large Reynolds numbers.

2. In the case of finite damping, the solution $\phi(y)$ of the characteristic-value problem does *not* tend to a solution of the inviscid equation along the whole section of the real axis of the complex y-plane which corresponds to the field of flow. (It does in a certain region of the complex y-plane.) The effect of viscosity persists at least at two layers in the interior of the fluid as the Reynolds numbers tend to infinity. The only possible exception to this statement is the singular case with $w''(y_c) = 0$.

3. As a corollary, it is concluded that damped and amplified disturbances are not complex conjugates in the limiting case, as shown by an examination of the inviscid equation (10).

These conclusions appear rather surprising at first sight, but they are confirmed by the work of Wasow [8]. Wasow further showed that the whole *region* between the two viscous layers in the damped case must be regarded as viscous.

The precise mathematical formulation and proof of the above conclusions will be discussed in the next section. We shall outline here a simple argument to support the conclusions.

The inviscid solution $\phi_2^{(0)}$ is multiple-valued (except when $w''(y_c) = 0$). We know that the correct branch is obtained by following a path *below* the

point y_c. The real axis, in the case of damped oscillations, is *above* y_c. Thus, the inviscid solution cannot be a valid approximation along the whole section of the real axis, and viscous effect must be present. Similar arguments show that, in the case of amplified solutions, the inviscid solution is a valid approximation along the real axis. This is essentially the original reasoning used by the present writer.

In the following two sections, the limiting process $\alpha R \to \infty$ will be discussed in some detail. It will be seen that there are several cases to be distinguished depending on the limiting behavior of the complex wave velocity c.

10. Cases in which the critical point does not approach the end points in the limit. In the general study of the inviscid problem, the critical point y_c, where $w = c$, does not coincide with either of the end points y_1 and y_2 where boundary conditions are imposed. If y_c does coincide with either y_1 or y_2, a special investigation is necessary. We shall therefore first consider the asymptotic approach of the viscous solution to those inviscid solutions for which $y_1 - y_c$ and $y_2 - y_c$ remain finite while $\alpha R \to \infty$.

In such cases, the characteristic-value problem can be studied by means of the asymptotic solutions of the types given by equations (9) and (11). It has been found by Wasow that the curves $Rl(Q) =$ constant, especially $Rl(Q) = 0$, are important for the asymptotic theory. The geometry of such curves relative to the real axis is shown in Fig. 4 for neutral, amplified, and damped disturbances (see also Fig. 2).

In this figure, the curves C_1, C_2, and C_3 are the three branches of the curve

$$Rl(Q) = 0.$$

The relative position of the real axis is also shown for all three cases. In the neutral case, it passes through the point y_c, where $w = c$. In the amplified case, the point y_c is above the real axis by an amount I which is proportional to c_i for small c_i. In the damped case, the point y_c is below the real axis by a corresponding amount. The points y_1 and y_2 on the real axis are the end points where the boundary conditions must be satisfied, with y_1 corresponding to the wall.

Clearly, other kinds of geometrical relations are possible. To fix our ideas, we shall limit ourselves to the kind shown in Fig. 4. The discussion of other possibilities will be taken up later.

The shaded areas represent regions where the solution $\phi_2^{(0)}$ of the inviscid equation (10) ceases to be a good asymptotic representation of a solution of the complete equation (10). In fact, the solution ϕ_2 which has the behavior

$$\phi_2 \sim \phi_2^{(0)} = 1 + a(y - y_c)^2 + \cdots + \frac{w_c''}{w_c'} \log (y - y_c)\phi_1$$

in S_1 and S_2 diverges [8] exponentially in the sector S_3,

$$\phi_2 \sim (\alpha R)^{-\frac{3}{4}} \exp (\sqrt{\alpha R}\, Q)[f_0 + O(\alpha R)^{-\frac{1}{2}}]$$

where Q is defined by (18) with $Rl(Q) > 0$ in S_3. The circular shaded region is associated with the fact that $d\phi_2^{(0)}/dy$ becomes infinite as $y - y_c \to 0$. From the solution of the form (16), it may be surmised that the radius of the shaded region is of the order of $(\alpha R)^{-\frac{1}{3}}$. This may be confirmed by the work of Wasow (cf. Sec. 4).

These shaded areas, related to the unusual behavior of the solution ϕ_2, are to be associated with the frictional effects in the interior of the fluid. The

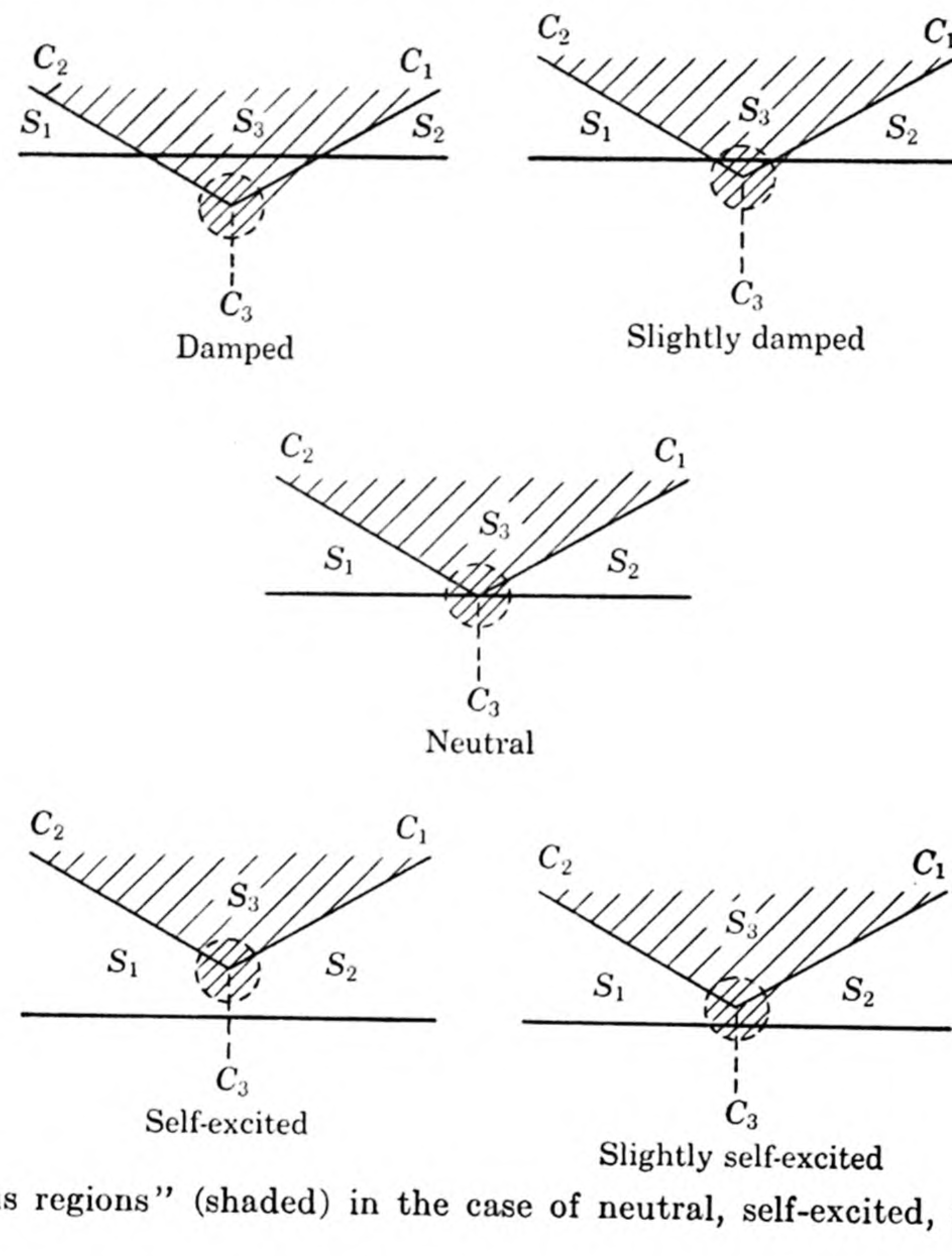


FIG. 4. "Viscous regions" (shaded) in the case of neutral, self-excited, and damped disturbances.

friction layer near the solid boundary—the outer friction layer—is given essentially correctly by the simple arguments in Sec. 9.

A description of inner and outer friction layers and regions can immediately be obtained from Fig. 4:

1. The outer and inner friction layers or regions are distinctly separated from each other (for αR sufficiently large).

2. The outer layer has a thickness of the order of magnitude of $(\alpha R)^{-\frac{1}{3}}$, approaching zero as $R \to \infty$.

3. In the case of finite amplification, there is no inner viscous layer for sufficiently large Reynolds numbers $[(\alpha R)^{-\frac{1}{3}} \ll c_i]$.

4. In the case of finite damping, there is a viscous *region* of finite width, however large the value of αR may become.

5. In the neutral case, the thickness of viscous layer is of the order of $(\alpha R)^{-\frac{1}{3}}$ and approaches zero as $\alpha R \to \infty$.

6. In the cases of damped and amplified disturbances with $|c_i| = O[(\alpha R)^{-\frac{1}{3}}]$—and hence approaching zero with $\alpha R \to \infty$—the disturbance has a behavior similar to the neutral case.

A few comments would be proper here in view of the earlier controversy which has centered around the above conclusions. First of all, it should be realized that the viscous effect is to be associated with the whole third sector S_3, besides the small circular region around y_c, which becomes vanishingly small as $\alpha R \to \infty$. Next, in the physical interpretation, one should distinguish between a viscous *layer*, whose thickness approaches zero as $\alpha R \to \infty$ (like the boundary layer of Prandtl), and a viscous *region* of finite width. This confusion is apparent in a comment by Holstein [24] on this subject. Holstein incorrectly stated that the present writer changed his conclusion [24] in asserting the existence of *one* viscous region. The existence of such a region does not alter the original conclusion of the present writer [4] that there *are* two viscous layers in the interior of the fluid in the limit of vanishing viscosity. It corroborates that statement with further information about the *region in between*.

It should also be noted that statement 6 does not help to answer the dilemma in the inviscid case, as raised in Sec. 9. In the limit of infinite Reynolds number, the statement concerns only the neutral case. These cases will be seen to have quite different characteristics. It appears that Holstein, without fully investigating the cases of finite damping and amplification, has used the results of the cases covered by statement 6 to contradict the author's discussion in the cases of finite damping and amplification.

This last point emphasizes the importance of distinguishing between different limiting processes. Here we have the difference between finite damping and damping approaching zero in the manner $c_i = O[(\alpha R)^{-\frac{1}{3}}]$. In the next section, we shall consider cases where $c \to w(y_1)$ or $w(y_2)$ as $\alpha R \to 0$.

11. Cases in which the critical point approaches one of the end points in the limit. If the critical point y_c does approach one of the end points y_1 or y_2, the situation is quite different. For definiteness, let us first consider the lower and upper branches of the neutral curve in the case of the plane Poiseuille flow.

(*a*) *Lower branch of the neutral curve.* Along this curve, the value of

$$\eta_1 = (y_1 - y_c)(\alpha R)^{\frac{1}{3}}$$

approaches a finite limit as $\alpha R \to \infty$. Thus the boundary point y_1 lies at a distance of the order of $\epsilon = (\alpha R)^{-\frac{1}{3}}$ from the critical point y_c, and therefore essentially within the circular viscous region around y_c (Fig. 4). There is no distinction between inner and outer friction layers. There is a *single* viscous layer which extends from the solid boundary to the critical point $y = y_c$ and *whose thickness approaches zero* as $\alpha R \to \infty$.

The same general conclusions hold for damped or amplified disturbances if the complex variable η_1 approaches a finite limit as $\alpha R \to \infty$. It should be

emphasized that in the limit of infinite Reynolds number the disturbance approaches a steady-state deviation with $c = 0$. There is no distinction between amplified, damped, or neutral oscillations in the limit.

However, it is possible for the characteristic value c to approach zero in a somewhat different manner, as happens in the next case.

(*b*) *Upper branch of the neutral curve.* Along the upper branch, the value of

$$\eta_1 = (y_1 - y_c)(\alpha R)^{\frac{1}{3}}$$

becomes infinite as $\alpha R \to \infty$. However, $y_1 - y_c$ still approaches zero. In fact, for plane Poiseuille motion,

$$\begin{aligned} \eta_1 &\sim (\alpha R)^{+\frac{2}{15}} \\ y_1 - y_c &\sim (\alpha R)^{-\frac{1}{5}} \end{aligned}$$

Note that η_1 is the ratio of the distance $y_1 - y_c$ to the thickness of the friction layer (radius of the circular region around the point y_c). Thus we may still regard the outer and inner friction layers as separated from each other.

The above are relatively simple cases. In the case of amplified and damped solutions, the condition $y_c \to y_1$ implies $c_r \to 0$ and $c_i \to 0$. The exact manner in which these two limiting processes depend on the Reynolds number αR may still vary, and this leads to descriptions of the inner and outer friction layers somewhat different from the above. We shall not go into the details of such discussions here.

12. Concluding remarks. The above discussions serve to illustrate the difficulties in the limiting process. It is necessary that these be completely understood; otherwise, one would wonder what happens in the limit of αR becoming infinite if one follows a path of constant α in Fig. 1. Obviously, the limiting point must correspond to a damped solution, but we know that the inviscid equation does not have a damped solution valid along the real axis. The above discussions show that the limit of a solution with finite damping does not satisfy the inviscid solution along the whole section of the real axis corresponding to the fluid field (except possibly in the exceptional case of $w''(y_c) = 0$), and consequently the effect of viscosity is not negligible in such limiting processes of vanishing viscosity.

Two important reasons may be suggested for the cause of the complication in the discussion of the limiting process.

1. Different limiting processes are often involved, with different results. This is clearly seen by contrasting the discussions of Secs. 10 and 11 and considering the cases in Sec. 11.

2. The solution ϕ_2 has an unusual behavior, as explained in Sec. 10. Without realizing this, one would have associated the viscous effects solely with the circular neighborhood of the size $(\alpha R)^{-\frac{1}{3}}$ around the critical point y_c. The difficulty in the formal limiting process, discussed in Sec. 9, would then appear.

Finally, it may be remarked that the behavior of the solution with finite damping (Sec. 10) closely resembles the structure of turbulence discussed by

Batchelor and Townsend [25]. The fluid motion exhibits a highly oscillatory behavior over a finite part of the field and very slow changes in another part. This again demonstrates an often stressed fundamental property of the motion of viscous fluids at high Reynolds numbers. In certain cases, the fluid behaves as a perfect fluid; in other cases, the effect of viscosity cannot be neglected even if it is very small. The increasingly refined spatial structure of the fluid motion is just enough to counterbalance the vanishing of viscosity and maintain the effects of the viscous terms in the Navier-Stokes equations of motion.

Note added in proof:

In a recent article, Tatsumi [26] critically discussed the method used by Pekeris and confirmed the work of the present writer. In two other articles [27], Tatsumi showed that the instability of the flow through a pipe could be caused by the instability of the laminar inlet flow prior to the formation of Poiseuille regime.

Bibliography

1. G. I. Taylor, Philos. Trans. Roy. Soc. London Ser. A vol. 223 (1923) pp. 289–343.
2. D. Meksyn, Proc. Roy. Soc. London Ser. A vol 187 (1946) pp. 115–128, 480–491, 492–504.
3. W. Heisenberg, Ann. d. Phys. vol. 74 (1924) pp. 577–627; Proc. Int. Congr. Math. vol. 2 (1950) pp. 292–296.
4. C. C. Lin, Quart. Appl. Math. vol. 3 (1945) pp. 117–142, 218–234, 277–301.
5. C. L. Pekeris, Phys. Rev. vol. 74 (1948) pp. 191–199.
6. L. H. Thomas, Phys. Rev. vol. 86 (1952) pp. 812–813.
7. H. B. Squire, Proc. Roy. Soc. London Ser. A vol. 142 (1933) pp. 621–628.
8. W. Wasow, Ann. of Math. vol. 58 (1953) pp. 222–252.
9. W. Tollmien, Zeit. Angew. Math. Mech. vols. 25–27 (1947) pp. 33–50, 70–83.
10. ———, Nachr. Ges. Wiss. Göttingen (1929) pp. 21–44.
11. D. W. Dunn and C. C. Lin, J. Aero. Sci. vol. 19 (1952) p. 491.
12. L. Lees and C. C. Lin, NACA Tech. Note 1115 (1946).
13. W. R. Dean, Proc. Roy. Soc. London Ser. A vol. 121 (1928) pp. 402–420.
14. H. Schlichting, Nachr. Ges. Wiss. Göttingen vol. 1 (1935) pp. 47–78.
15. G. B. Schubauer and H. K. Skramstad, J. Aero. Sci. vol. 14 (1947) pp. 69–78.
16. L. Lees, NACA Tech. Note 1360 (1946), reissued as NACA Tech. Rep. 876 (1947).
17. P. Chiarulli, Ph.D. Thesis, Brown University, 1949.
18. M. Lessen, NACA Tech. Note 1929 (1947), reissued as NACA Tech. Rep. 979 (1950).
19. S. I. Pai, J. Aero. Sci. vol. 18 (1951) pp. 731–742.
20. C. C. Lin, NACA Tech. Note 2887 (1953).
21. J. R. Foote and C. C. Lin, Quart. Appl. Math. vol. 8 (1951) pp. 265–280.
22. C. G. Rossby, J. Meteor. vol. 6 (1949) pp. 163–180.
23. H. L. Kuo, Tellus vol. 3 (1951) pp. 268–284.
24. Letter of W. Wasow and C. C. Lin, and H. Holstein's comment, Zeit. Angew. Math. Mech. vol. 30 (1950) pp. 25–49; vol. 31 (1951) p. 160.
25. G. K. Batchelor and A. A. Townsend, Proc. Roy. Soc. London Ser. A vol. 199 (1949) pp. 238–255.
26. T. Tatsumi, J. Phys. Soc. Japan vol. 7 (1952) pp. 619–624.
27. ———, J. Phys. Soc. Japan vol. 7 (1952) pp. 489–495, 495–502.

Massachusetts Institute of Technology,
Cambridge, Mass.

EXAMPLES OF THE INSTABILITY OF FLUID MOTION IN THE PRESENCE OF A MAGNETIC FIELD

BY

S. CHANDRASEKHAR

1. Introduction. While the stability of two-dimensional plane flows has only recently been settled beyond dispute (cf. Lin [1]), there have been two other examples of hydrodynamic stability which though they have attracted much less attention have been fully understood and are probably of even greater practical significance. The two examples are the thermal instability of a horizontal layer of fluid heated below and the rotational instability of viscous flow between rotating cylinders. In this paper we shall reconsider these two classical problems in the framework of hydromagnetics, *i.e.*, when the fluid considered is an electrical conductor and a magnetic field is present. But we shall first summarize the known results in hydrodynamics.

2. Thermal instability in hydrodynamics. The manner of the onset of instability in a horizontal layer of incompressible fluid has been the subject of investigations by Rayleigh [2], Jeffreys [3], Low [4], Pellew and Southwell [5], and others. The principal results established by these investigations are: (i) A layer of fluid heated below first becomes unstable when the Rayleigh number

$$R = \frac{g\alpha\beta}{\kappa\nu} d^4 \tag{1}$$

(where g denotes the value of gravity, d the depth of the layer considered, $-\beta$ is the adverse temperature gradient which is maintained, and α, κ, and ν are the coefficients of volume expansion, thermometric conductivity, and kinematic viscosity, respectively) exceeds a certain determinate critical value. (ii) The motions which ensue on surpassing the critical value must have a cellular pattern in agreement with what had been demonstrated experimentally by Bénard [6]. The exact value of the Rayleigh number at which instability sets in depends on boundary conditions such as whether a bounding surface is free or rigid. More particularly, the characteristic-value problem, the solution of which leads to the determination of the critical Rayleigh number, is the following:

For a given a^2 (>0) find the smallest characteristic value R for which the equation

$$(D^2 - a^2)^3 W = -a^2 R W \qquad \left[D = \frac{d}{dz}, W = W(z)\right] \tag{2}$$

admits a solution which satisfies the boundary conditions

$$W = (D^2 - a^2)^2 W = 0 \qquad (\text{for } z = \pm\tfrac{1}{2}) \tag{3}$$

and

$$\text{(4)} \qquad \textit{either} \quad D^2W \quad \textit{or} \quad DW = 0,$$

depending on whether the bounding plane at $z = \frac{1}{2}$ (similarly $z = -\frac{1}{2}$) is free or rigid. Then determine the minimum of this lowest characteristic value as a function of a^2; this minimum will specify the critical Rayleigh number R_c at which convection through instability will first set in.

From a direct solution of the foregoing problem it has been derived that (cf. Pellew and Southwell [5])

$$\text{(5)} \qquad \begin{aligned} R_c &= 657.51, && \text{(both bounding surfaces free)},\\ R_c &= 1{,}100.6, && \text{(one bounding surface free and the other rigid)},\\ R_c &= 1{,}707.8, && \text{(both bounding surfaces rigid)}. \end{aligned}$$

It may be noted here that the onset of instability by convection predicted at $R = 1{,}707.8$ when the layer of fluid is confined between two rigid planes has been confirmed experimentally by Schmidt and Milverton [7].

3. Rotational instability in hydrodynamics. The stability of viscous flow between two concentric rotating cylinders was first successfully treated both experimentally and theoretically by G. I. Taylor [8]. (Later investigations of the same problem are those of Jeffreys [3] and Meksyn [9].) In this case the problem is the following: The hydrodynamical equations allow the stationary solution

$$\text{(6)} \qquad V(r) = Ar + \frac{B}{r}$$

for the rotational velocity at a distance r from the axis, where A and B are two constants related to the angular velocities Ω_1 and Ω_2 with which the inner and the outer cylinders (of radii R_1 and R_2) are rotated. Thus

$$\text{(7)} \qquad A = \Omega_1 \frac{1 - mR_2^2/R_1^2}{1 - R_2^2/R_1^2} \quad \text{and} \quad B = \Omega_1 \frac{R_1^2(1 - m)}{1 - R_1^2/R_2^2},$$

where $m = \Omega_2/\Omega_1$. By considering a symmetrical perturbation of solution (6) by a disturbance characterized by a wave number λ, it can be shown that the situation in marginal stability is governed by the equation

$$\text{(8)} \qquad (DD^* - \lambda^2)^3 v = \frac{4A\lambda^2}{\nu^2}\left(A + \frac{B}{r^2}\right)v,$$

where

$$\text{(9)} \qquad D = \frac{d}{dr} \quad \text{and} \quad D^* = D + \frac{1}{r},$$

together with the boundary conditions

$$\text{(10)} \qquad v, \quad (DD^* - \lambda^2)v, \quad \text{and} \quad D^*(DD^* - \lambda^2)v$$

all vanish for $r = R_1$ and $r = R_2$.

The characteristic-value problem presented by equation (8) and boundary conditions (10) has not been treated in its full generality. The only circumstance under which the problem has proved amenable to treatment is when the difference in radii of the two cylinders is small compared with their mean, *i.e.*, when

$$d = R_2 - R_1 \ll \tfrac{1}{2}(R_2 + R_1) = R_0. \tag{11}$$

When (11) holds, we need not distinguish between D and D^* (cf. Meksyn [9]). And if we further restrict ourselves to the case when the two cylinders are rotating in the same direction, an additional simplification is possible in that we may replace $A + (B/r^2)$ in (8) by the value Ω_0 of the angular velocity at $r = R_0$. Under these circumstances the characteristic value problem reduces to solving

$$(D^2 - a^2)^3 v = -a^2 T v \tag{12}$$

with the boundary conditions

$$v = (D^2 - a^2)v = D(D^2 - a^2)v = 0, \qquad (\text{for } z = \pm\tfrac{1}{2}). \tag{13}$$

In equation (12)

$$T = -\frac{4\Omega_0 A}{\nu^2} d^4 \tag{14}$$

is the Taylor number. It can be readily shown that this characteristic-value problem is identical with that appropriate for the problem of thermal instability in the case when both the bounding surfaces are rigid. Hence, the critical Taylor number for instability is [cf. equation (5)]

$$T = 1{,}707.8. \tag{15}$$

As was first pointed out by Low [4], this equality between the Rayleigh and the Taylor numbers has a simple physical origin. And it may also be noted here that the theoretical value (15) is in excellent agreement with the experimental results of G. I. Taylor.

4. The equations of hydromagnetics. If the fluid considered is an electrical conductor and a magnetic field is present, the relevant equations are

$$\rho \frac{\partial \mathbf{u}}{\partial t} + \rho(\mathbf{u} \cdot \operatorname{grad})\mathbf{u} - \mu \mathbf{J} \times \mathbf{H} = -\operatorname{grad} p + \rho \operatorname{grad} V + \rho\nu\nabla^2\mathbf{u}, \tag{16}$$

$$\operatorname{curl} \mathbf{H} = 4\pi\mathbf{J}, \qquad \operatorname{curl} \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}, \tag{17}$$

$$\operatorname{div} \mathbf{H} = 0, \qquad \mathbf{J} = \sigma(\mathbf{E} + \mu\mathbf{u} \times \mathbf{H}), \tag{18}$$

and

$$\frac{\partial T}{\partial t} + (\mathbf{u} \cdot \operatorname{grad})T = \kappa\nabla^2 T \tag{19}$$

where ρ is the density, p the pressure, T the temperature, V the gravitational potential, $\mathbf{E}$ the intensity of the electric field, $\mathbf{H}$ the intensity of the magnetic

field, **J** the current density, μ the magnetic permeability, and σ the coefficient of electrical conductivity. By eliminating **E** and **J** we can combine equations (16) to (18) to give the following pair of equations:

$$(20)\qquad \rho\frac{\partial u_i}{\partial t} + \rho u_j\frac{\partial u_i}{\partial x_j} - \frac{\mu}{4\pi}H_j\frac{\partial H_i}{\partial x_j} = -\frac{\partial}{\partial x_i}\left(p + \mu\frac{|\mathbf{H}|^2}{8\pi}\right) + \rho\nu\nabla^2 u_i + \rho\frac{\partial V}{\partial x_i}$$

and

$$(21)\qquad \frac{\partial H_i}{\partial t} + u_j\frac{\partial H_i}{\partial x_j} - H_j\frac{\partial u_i}{\partial x_j} = \eta\nabla^2 H_i,$$

where

$$(22)\qquad \eta = \frac{1}{4\pi\mu\sigma}.$$

Equations (19), (20), and (21) together with div $\mathbf{H} = 0$ and the equation of continuity are the basic equations of hydromagnetics.

5. The problem of thermal instability in hydromagnetics. Restricting ourselves to the case when the impressed magnetic field (H) and gravity act in the same direction, we find that the critical Rayleigh number for the onset of instability depends on the strength of the magnetic field and the electrical conductivity, σ, through the nondimensional parameter

$$(23)\qquad Q = \frac{\mu^2 H^2\sigma}{\rho\nu}d^2.$$

Indeed, we find that the characteristic-value problem which we have to solve for determining the critical Rayleigh number is [cf. equations (2) to (4)]

$$(24)\qquad (D^2 - a^2)[(D^2 - a^2)^2 - QD^2]W = -a^2RW$$

together with the boundary conditions

$$(25)\qquad W = [(D^2 - a^2)^2 - QD^2]W = 0 \qquad (\text{for } z = \pm\tfrac{1}{2})$$

and

$$(26)\qquad \textit{either}\quad D^2W \quad \textit{or} \quad DW = 0,$$

depending on whether the bounding plane at $z = \frac{1}{2}$ (similarly $z = -\frac{1}{2}$) is free or rigid. And as in the case of zero Q we must determine the minimum of the lowest characteristic value R of the foregoing problem as a function of a^2.

It has been found (Chandrasekhar [10]) that the most convenient method of solving the problem presented by equations (24) to (26) is by making use of the following variational principle: The formula

$$(27)\qquad R = \frac{\int_{-\frac{1}{2}}^{+\frac{1}{2}}[(DF)^2 + a^2F^2]\,dz}{a^2\int_{-\frac{1}{2}}^{+\frac{1}{2}}\{[(D^2 - a^2)W]^2 + Q(DW)^2\}\,dz},$$

where

$$(28)\qquad F = [(D^2 - a^2)^2 - QD^2]W,$$

provides a minimum value of R if the boundary conditions of the problem are satisfied and

$$(D^2 - a^2)F = -a^2RW \tag{29}$$

(*i.e.*, if the equation governing W is satisfied). The proof of this variational principle will be found in Chandrasekhar [10]. But here it will suffice to note that it enables the solution of equations (24) to (26) by the following variational procedure:

Assume for F an expression involving one or more parameters A_k and which vanishes for $z = \pm\frac{1}{2}$. With the chosen form of F determine W as a solution of the equation

$$[(D^2 - a^2)^2 - QD^2]W = F \tag{30}$$

and satisfying the boundary conditions on W at $z = \pm\frac{1}{2}$. Then evaluate R according to formula (27) and minimize it with respect to the parameters A_k. In this way we shall obtain the "best" value of R for the chosen form of F.

The critical Rayleigh numbers listed in Table I are an extract from a more extensive tabulation given in [10].

TABLE I

Q	R_c		
	Two free boundaries	Two rigid boundaries	One free boundary One rigid boundary
0	657.5	1,708	1,101
100	2,654	3,757	1,699
500	8,579	10,110	3,586
1,000	15,210	17,100	5,613
10,000	119,800	124,500	35,040

In order to see how effective the magnetic field will be in practical cases in inhibiting the onset of convection, we shall consider mercury at room temperature. We have $\sigma = 1.1 \times 10^{-5}$, $\rho\nu = 1.7 \times 10^{-2}$, and $Q = 6 \times 10^{-4}H^2d^2$, where H is measured in gauss. For $d = 1$ cm. and $H = 1{,}000$ gauss, $Q = 600$, and the critical Rayleigh numbers for the three cases considered are

$$\begin{aligned} R_c &= 9{,}942, & \frac{R_c}{R_0} &= 15.12, & &[\text{case (i)}],\\ R_c &= 11{,}660, & \frac{R_c}{R_0} &= 6.80, & &[\text{case (ii)}],\\ R_c &= 4{,}018, & \frac{R_c}{R_0} &= 3.65, & &[\text{case (iii)}]. \end{aligned} \tag{31}$$

It would appear from these values that the predicted effect should be easily detectable in the laboratory.

6. The problem of rotational instability in hydromagnetics. Turning next to the problem of the stability of viscous flow between rotating cylinders in hydromagnetics, we shall consider the case when the external magnetic field is in the z-direction, *i.e.*, along the axis of the cylinders. In this case the equations of the problem (cf. Sec. 4) admit the same stationary solution (6) as when no field is present. And considering symmetrical perturbations of this solution which are periodic in z and restricting ourselves to the case when $d \ll R_0$ [cf. equation (11)] and the two cylinders are rotating in the same direction, we find that the characteristic-value problem, the solution of which leads to the determination of the critical Taylor number for instability, is (Chandrasekhar [11])

$$[(D^2 - a^2)^2 + Qa^2]^2\psi = -a^2T(D^2 - a^2)\psi \tag{32}$$

together with the boundary conditions

$$D\psi = (D^2 - a^2)\psi = [(D^2 - a^2)^2 + Qa^2]\psi = D[(D^2 - a^2)^2 + Qa^2]\psi = 0, \quad (\text{for } z = \pm\tfrac{1}{2}), \tag{33}$$

where Q has the same meaning as in equation (23). Again as in the case of zero Q we must determine the minimum of the lowest characteristic value of the foregoing problem as a function of a^2.

It has been found (Chandrasekhar [11]) that the most convenient method of solving the problem presented by equations (32) and (33) is by appealing to the following variational principle: The formula

$$T = \frac{\int_{-\frac{1}{2}}^{+\frac{1}{2}} \{[(D^2 - a^2)P]^2 + Qa^2P^2\}\, dz}{a^2 \int_{-\frac{1}{2}}^{+\frac{1}{2}} \{(DG)^2 + a^2G^2 + a^2Q[(D\psi)^2 + a^2\psi^2]\}\, dz}, \tag{34}$$

where

$$P = [(D^2 - a^2)^2 + Qa^2]\psi \quad \text{and} \quad G = (D^2 - a^2)\psi, \tag{35}$$

provides a minimum value of T if the boundary conditions of the problem are satisfied and

$$[(D^2 - a^2)^2 + Qa^2]P + a^2TG = 0 \tag{36}$$

(*i.e.*, if the equation governing ψ is satisfied). The proof of this variational principle will be found in Chandrasekhar [11]. But here it will suffice to note that it enables the solution of equations (32) and (33) by the following variational procedure:

Assume for DP an expression involving one or more parameters A_k and which vanishes for $z = \pm\frac{1}{2}$. Integrating the expression for DP, obtain P, and adjust the constant of integration so that P may also vanish for $z = \pm\frac{1}{2}$. (The fact that both P and DP must vanish for $z = \pm\frac{1}{2}$ implies a certain restriction in the form for DP we can assume.) Next solve the equation

$$[(D^2 - a^2)^2 + Qa^2]\psi = P \tag{37}$$

for ψ, and determine the constant of integration so as to satisfy the boundary conditions $D\psi = (D^2 - a^2)\psi = 0$ for $z = \pm\frac{1}{2}$. And finally evaluate T according to the formula (34), and minimize it with respect to the parameters A_k. In this way we shall obtain the "best" value of T for the chosen form of DP.

The critical Taylor numbers listed in Table II are an extract from a more extensive tabulation given in [11].

TABLE II

Q	T_c	Q	T_c
0	1,708	1,000	378,600
10	2,691	10,000	4,459,000
100	17,570		

In order to see how effective the magnetic field will be in practical cases in inhibiting rotational instability, we shall again consider mercury at room temperature. Using the values of the physical constants appropriate for mercury, we find that, for $d = 1$ cm. and $H = 1{,}000$ gauss, Q has the value 600; the corresponding critical Taylor number is 80 times the value for zero field. The predicted effect is therefore very pronounced and should be detectable in the laboratory. Indeed, it would appear that the effect should be detectable even with ordinary acids and acid solutions. Thus for nitric acid at room temperature $\sigma = 8 \times 10^{-10}$, $\rho\nu = 2 \times 10^{-2}$, and $Q = 4 \times 10^{-8}H^2d^2$. Accordingly, under ordinary laboratory conditions the effect should become measurable for $H \sim 10^4$ gauss.

Finally, we may compare the inhibiting effect of a magnetic field on rotational and thermal instabilities. As we have already pointed out, in the absence of a magnetic field, the critical Rayleigh number for thermal instability of a layer of fluid confined between two rigid planes is the same as the critical Taylor number for rotational instability under the circumstances in which the problem has been investigated (*i.e.*, when the cylinders are rotating in the same direction and $d \ll R_0$). When a magnetic field is present, this identity between the Taylor and the Rayleigh numbers no longer exists though they both continue to be functions of the same nondimensional parameter Q. Also it is seen that the inhibition of rotational instability by a magnetic field is very much more pronounced than the inhibition of thermal instability.

7. Some related problems in hydrodynamic and hydromagnetic stability. The foregoing discussion of the classical problems of thermal instability and rotational instability has not included any reference to a number of related problems which have recently been solved. We may therefore conclude by briefly describing these further developments of the theory.

(i) The problem of the thermal instability of a layer of fluid heated below has been extended (Chandrasekhar [12]) to include the effect of Coriolis forces. And it has been found that when Coriolis forces are included the critical Ray-

leigh number for the onset of instability depends on the angular velocity of rotation Ω through the nondimensional parameter

$$T = \frac{4\Omega^2}{\nu^2} d^4 \tag{38}$$

in such a way that

$$R \to \text{const. } T^{2/3} \quad \text{as} \quad T \to \infty. \tag{39}$$

The changed dependence of β on ν and d which this entails has bearing on a number of meteorological and astrophysical questions.

(ii) The problem of the thermal instability of a layer of fluid heated below and subject simultaneously to Coriolis forces and an external magnetic field has also been investigated (Chandrasekhar [13]), and the results exhibit in a striking way the conflicting tendencies to which a conducting fluid is subject in its attempt to attach itself simultaneously to the vortex lines and the lines of magnetic force.

(iii) The problem of the thermal instability of a fluid sphere heated within has been treated by Jeffreys and Bland [14] and Chandrasekhar [15]. The analysis shows that the pattern of convection which is easiest to excite is a spherical harmonic of order 1.

(iv) The thermal instability of an incompressible sphere consisting of an inviscid core and a viscous mantle has also been considered (Chandrasekhar [16]), and it has been shown that the pattern of convection which sets in, at marginal stability, in the mantle shifts to harmonics of the higher orders as the thickness of the mantle decreases. Thus, when the mantle extends to a depth of half the radius of the sphere, the harmonics of order 3 and 4 set in about simultaneously, while the harmonic of order 5 follows very soon afterward. These results have a bearing on the interpretation of the earth's topographic features (cf. Vening Meinesz [17]).

Finally it may be stated that the solution of the various problems listed above was possible only by suitable extensions of the variational methods described in Secs. 5 and 6.

Bibliography

1. C. C. Lin (the preceding paper in this volume).

2. Lord Rayleigh, *On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side*, in *Scientific papers*, Vol. 6 Cambridge University Press, London, 1920, pp. 432–446.

3. H. Jeffreys, *The stability of a layer of fluid heated below*, Philos. Mag. Ser. 7 vol. 2 (1926) pp. 833–844.

———, *Some cases of instability in fluid motion*, Proc. Roy. Soc. London Ser. A vol. 118 (1928) pp. 195–208.

4. A. R. Low, *On the criterion for stability of a layer of viscous fluid heated from below*, Proc. Roy. Soc. London Ser. A vol. 125 (1929) pp. 180–195.

5. A Pellew and R. V. Southwell, *On maintained convective motion in a fluid heated from below*, Proc. Roy. Soc. London Ser. A vol. 176 (1940) pp. 312–343.

6. H. Bénard, *Les tourbillons cellulaires dans une nappe liquide transportant de la chaleur par convection en régime permanent*, Ann. Chim. (Phys.) vol. 23 (1901) pp. 62–144.

7. R. J. Schmidt and S. W. Milverton, *On the stability of a fluid when heated from below*, Proc. Roy. Soc. London Ser. A vol. 152 (1935) pp. 586–594.

8. G. I. Taylor, *Stability of a viscous liquid contained between two rotating cylinders*, Philos. Trans. Roy. Soc. London Ser. A vol. 223 (1923) pp. 289–343.

9. D. Meksyn, *Stability of viscous flow between rotating cylinders*, I, II, III, Proc. Roy. Soc. London Ser. A vol. 187 (1946) pp. 115–128, 480–491, 492–504.

10. S. Chandrasekhar, *On the inhibition of convection by a magnetic field*, Philos. Mag. Ser. 7 vol. 43 (1952) pp. 501–532.

11. ———, *The stability of viscous flow between rotating cylinders in the presence of a magnetic field*, Proc. Roy. Soc. London Ser. A vol. 216 (1953) pp. 293–309.

12. ———, *The instability of a layer of fluid heated below and subject to Coriolis forces*, Proc. Roy. Soc. London Ser. A vol. 217 (1953) pp. 306–327.

13. ——— (in press).

14. H. Jeffreys and M. E. M. Bland, *The instability of a fluid sphere heated within*, Mon. Not. R. Astr. Soc. Geophys. Sup. vol. 6 (1951) pp. 148–158.

15. S. Chandrasekhar, *The thermal instability of a fluid sphere heated within*, Philos. Mag. Ser. 7 vol. 43 (1953) pp. 1317–1329.

16. ———, *The onset of convection by thermal instability in spherical shells*, Philos. Mag. Ser. 7 vol. 44 (1953) pp. 233–241, 1129–1130.

17. F. A. Vening Meinesz, *A remarkable feature of the earth's topography, origin of continents and oceans*, I, II, Neder. Akad. Wetensch. Ser. B vol. 54 (1951) pp. 212–219, 220–228.

Yerkes Observatory, University of Chicago,
Williams Bay, Wis.

ON FREE-SURFACE FLOWS

BY

P. R. GARABEDIAN

1. The principle of minimum virtual mass. It is the purpose of this paper to review briefly some researches on axially symmetric cavitational flow which have been carried out jointly by H. Lewy, M. Schiffer, and the present author at the Applied Mathematics and Statistics Laboratory of Stanford University.

The study of axially symmetric cavity flow has acquired particular interest with the advent of modern projectiles moving rapidly through water. The physical problem is characterized by motion of a projectile so rapid that the fluid pressure falls below the vapor pressure of water. Thus a cavity of steam of unknown shape forms behind the projectile. Calculation of the shape of the cavity is the mathematical problem to which this physical phenomenon leads.

We study here the axially symmetric motion of an ideal incompressible fluid of density 1 past an axially symmetric projectile B in a cylindrical tunnel. We assume that the motion is steady and irrotational. It can be represented in terms of a velocity potential φ or a stream function ψ. Let the x-axis be the axis of symmetry and let the xy-plane be a meridian cross section of space. We can represent φ and ψ as functions of x,y alone, and they satisfy the pair of generalized Cauchy-Riemann equations

$$\varphi_x = \frac{\psi_y}{y}, \qquad \varphi_y = -\frac{\psi_x}{y}. \tag{1}$$

It is sufficient, by the above symmetry, to investigate φ and ψ for nonnegative values of y, and hence we shall consider only the half space $y \geqq 0$ in the following.

We require that the object B have a surface generated by rotation about the x-axis of a curve C which is monotonically increasing for $x < 0$ and monotonically decreasing for $x > 0$ and is symmetric in the y-axis. We shall not distinguish in our notation between a given geometrical figure in space and its meridian cross section in the upper half of the xy-plane. The curve C is assumed to intersect the x-axis at $(-k_0,0)$ and $(k_0,0)$, and it is assumed to consist within the interval $-k \leqq x \leqq k$, $k_0 > k > 0$, of a horizontal segment $y = h$, $h > 0$. The case where C rises vertically from the x-axis is allowed. We require that the object B lie in the tunnel $0 \leqq y < h_0$, in which the flow is supposed to take place, whence it is necessary to assume that $h < h_0$.

We denote by L a curve, or set of curves, which bounds together with the segment $y = h$, $-k \leqq x \leqq k$, a cavity region W which lies outside B, but whose closure is contained within the tunnel $0 \leqq y < h$ and within the strip $-k \leqq x \leqq k$. We denote by D the region of the tunnel $0 \leqq y < h$ which is comple-

mentary to $B + W$. It is within the region D that the above axially symmetric flow of water takes place. B is visualized as the projectile, whereas W will be taken to represent a cavity of steam.

By elimination of φ from (1), we see that ψ must satisfy in the flow region D the second-order elliptic partial differential equation

$$\psi_{xx} + \psi_{yy} = \frac{\psi_y}{y}. \tag{2}$$

As boundary conditions on ψ we require that along the tunnel wall $y = h_0$

$$\psi = \frac{h_0^2}{2}, \tag{3}$$

while on those arcs of C and L which bound D we require

$$\psi = 0. \tag{4}$$

On the intervals $x < -k_0$, $k_0 < x$, of the x-axis ψ is regular and vanishes. We suppose that ψ and its derivatives are bounded at infinity.

The virtual mass M of the axially symmetric flow ψ is defined by the integral

$$M = 2\pi \iint_D \left(\nabla\psi - \nabla \frac{y^2}{2}\right)^2 \frac{dx\, dy}{y}, \tag{5}$$

since $\psi - (y^2/2)$ represents the stream function of an unsteady flow in which $B + W$ moves with uniform velocity against a fluid at rest at the infinite ends of the tunnel $0 \leqq y < h_0$. We denote by

$$V = 2\pi \iint_{B+W} y\, dx\, dy \tag{6}$$

the volume of the projectile B and cavity W.

These preliminaries enable us to set up an extremal problem which will characterize the axially symmetric flow ψ as a cavitational flow. We suppose that the projectile B is fixed, whereas the cavity W is allowed to have any shape lying within the rectangle $-k \leqq x \leqq k$, $h \leqq y < h_0$. For each fixed $\mu > 0$, we seek to choose the cavity W in such a way that

$$M - \mu V = \text{minimum}. \tag{7}$$

The extremal problem (7) serves to determine the free-boundary curve L of W.

Let us suppose for the moment that an analytic curve L exists bounding an extremal shape W for (7). We suppose that a small shift of L is made by δn units along the inner normal of L with respect to D. This shift generates a varied curve L^* bounding a varied cavity W^* and a varied flow region D^* with stream function ψ^*. We denote by M^* and V^* the virtual mass of D^* and the volume of $B + W^*$. It is not too difficult to show [2, 10] that within

accuracy of the first order in δn

(8) $$\delta V = V^* - V = 2\pi \int_L y\, \delta n\, ds,$$

(9) $$\delta M = M^* - M = 2\pi \int_L \left[\frac{1}{y^2}\left(\frac{\partial\psi}{\partial n}\right)^2 - 1\right] y\, \delta n\, ds,$$

where s is the arc length along L. But in virtue of (7)

(10) $$\delta M - \mu\, \delta V = 2\pi \int_L \left[\frac{1}{y^2}\left(\frac{\partial\psi}{\partial n}\right)^2 - 1 - \mu\right] y\, \delta n\, ds = 0$$

for every choice of δn. Hence along the extremal free boundary L we must have

(11) $$\frac{1}{y^2}\left(\frac{\partial\psi}{\partial n}\right)^2 = 1 + \mu.$$

The identity (11) shows that the velocity of the extremal flow is constant on L. On the other hand, Bernoulli's law gives for the pressure p the relation

(12) $$\frac{1}{2y^2}(\nabla\psi)^2 + p = \text{const.}$$

Thus on L, p has the constant value

$$p = p_0 - \frac{\mu}{2},$$

where p_0 is the pressure at infinity. This is precisely the condition on the above flow necessary in order that W be a cavity of steam at constant pressure. Therefore our extremal flow is a cavitating flow with free boundary L and cavitation parameter μ.

The earliest reference to the extremal problem (7) as a characterization of free-surface flow appears in the work of Riabouchinsky [10]. Existence theory for plane cavitational flow [2] and for axially symmetric cavitational flow [3] can be based on this principle. We shall proceed in the next few sections to indicate how a rigorous treatment of the extremal problem (7) can be carried through and thus how an existence theorem for axially symmetric cavitating flow can be obtained.

2. Existence of a rectifiable solution. A first analysis of (7) can be made using the notion of symmetrization [9].

Symmetrization of the object $B + W$ in the y-axis is defined to be the replacement in every meridian plane of the intersection of each horizontal line with $B + W$ by a single segment of that line which is bisected by the y-axis and has the same length as the linear measure of the original intersection. Symmetrization of $B + W$ in the x-axis shall mean in this paper replacement of the volume of revolution $B + W$ in space by a new volume with the property that its intersection with any plane perpendicular to the x-axis is a circular disk with area equal to the two-dimensional measure of the intersection of that

plane with the original volume. Symmetrization in the x-axis in space thus generates a certain modified symmetrization in the x-axis of the cross section of $B + W$ in the meridian xy-plane.

Because of the symmetry and monotonicity of the boundary curve C of B, we see that B is unchanged by symmetrization in the x-axis or y-axis. Also, it is clear that the volume V of $B + W$ is invariant under symmetrization of $B + W$. On the other hand, it can be shown from Schwarz's inequality or Minkowski's inequality that the virtual mass M of the axially symmetric flow ψ past $B + W$ is either diminished or left unchanged when $B + W$ is symmetrized in the x-axis or in the y-axis [2, 3, 8].

Thus a competing shape of W for the extremal problem (7) yields after symmetrization of $B + W$ an admissible new shape with the same value of V and a value of M which is either the same or smaller. Therefore it is sufficient to restrict the discussion to shapes such that $B + W$ is already symmetrized in the x-axis and in the y-axis. For a symmetrized shape of this type, the free curve L ascends monotonically in the second quadrant and descends monotonically in the first quadrant.

The virtual mass M of $B + W$ can be estimated in terms of the virtual mass of a circular disk perpendicular to the x-axis with a radius equal to the maximum height of L above the x-axis [3]. This estimate can be used to show that the quantity $M - \mu V$ in (7) increases indefinitely if L rises toward the tunnel wall $y = h_0$.

We conclude that an extremal shape for W satisfying (7) exists. For any symmetrized shapes W_n are bounded by monotonic curves L_n which satisfy a uniform Lipschitz condition in a coordinate system obtained by rotating the x-axis and y-axis through 45°. Hence, by equicontinuity, a convergent subsequence of shapes can be obtained from any sequence W_n for which the associated quantities $M_n - \mu V_n$ approach their greatest lower bound [3]. The limiting shape W is extremal for (7) and is bounded by a curve L which is monotonic in each quadrant. Furthermore, the level curves of the stream function ψ corresponding to W bound with the x-axis regions which are symmetrized, and hence these level curves ascend in the second quadrant and descend in the first quadrant.

3. Variational formulas. Let z_0 be any point in the interior of an arc of the extremal free boundary L and within the strip $-k < x < k$. Let ρ be a positive number so small that the circle $|z - z_0| \leqq 2\rho$ does not intersect B or the lines $y = h_0$, $x = -k$, $x = k$, and let ω be a function of $|z - z_0|$, with derivatives of all orders, which vanishes for $|z - z_0| \geqq 2\rho$ and is 1 for $|z - z_0| \leqq \rho$. For t in the exterior of D we set

$$F(z,\bar{z}) = \frac{\omega}{z - t}. \tag{13}$$

For t in the interior of D, we choose $\eta > 0$ so small that the circle $|z - t| \leqq \eta$ is in D, and we set

$$F(z,\bar{z}) = \frac{\omega}{z-t}, \qquad |z-t| > \eta, \tag{14}$$

$$F(z,\bar{z}) = \omega\,\frac{\bar{z}-\bar{t}}{\eta^2}, \qquad |z-t| \leqq \eta. \tag{15}$$

For sufficiently small values of the complex parameter ϵ, we consider the transformation

$$z^* = z + \epsilon F(z,\bar{z}) \tag{16}$$

of the independent variable $z = x + iy$. This transformation maps W onto a varied cavity W^* which is admissible for the external problem (7), and it maps D onto a varied flow region D^* with stream function ψ^*. We define further

$$\psi^{**}(x^*,y^*) = \psi(x,y), \tag{17}$$

with $z = x + iy$ and $z^* = x^* + iy^*$ related by (16). The functions ψ^{**} and ψ^* have the same values on the boundary of D^*, and hence by Dirichlet's principle

$$\iint_{D^*} \left(\nabla\psi^{**} - \nabla\frac{y^2}{2}\right)^2 \frac{dx\,dy}{y} \geqq \iint_{D^*} \left(\nabla\psi^{*} - \nabla\frac{y^2}{2}\right)^2 \frac{dx\,dy}{y}, \tag{18}$$

with $\psi^{**} = \psi^{**}(x,y)$, $\psi^* = \psi^*(x,y)$. On the other hand, by (7)

$$\iint_{D^*} \left(\nabla\psi^{*} - \nabla\frac{y^2}{2}\right)^2 \frac{dx\,dy}{y} - \mu \iint_{W^*} y\,dx\,dy \geqq \iint_{D} \left(\nabla\psi - \nabla\frac{y^2}{2}\right)^2 \frac{dx\,dy}{y} - \mu \iint_{W} y\,dx\,dy. \tag{19}$$

Combining (18) and (19), we obtain

$$\iint_{D^*} \left(\nabla\psi^{**} - \nabla\frac{y^2}{2}\right)^2 \frac{dx\,dy}{y} - \mu \iint_{W^*} y\,dx\,dy \geqq \iint_{D} \left(\nabla\psi - \nabla\frac{y^2}{2}\right)^2 \frac{dx\,dy}{y} - \mu \iint_{W} y\,dx\,dy. \tag{20}$$

If we expand both sides of (20) in powers of ϵ and $\bar{\epsilon}$ and use the knowledge that ϵ can be chosen arbitrarily in a small neighborhood of the origin, we find by the usual technique of the calculus of variations

$$(\mu + 1) \iint_{W} \left(\frac{F}{i} + 2y\frac{\partial F}{\partial z}\right) dx\,dy + 4 \iint_{D} \frac{F}{i}\frac{\partial\psi}{\partial z}\frac{\partial\psi}{\partial\bar{z}}\frac{dx\,dy}{y^2} + 8 \iint_{D} \frac{\partial F}{\partial\bar{z}}\left(\frac{\partial\psi}{\partial z}\right)^2 \frac{dx\,dy}{y} = 0, \tag{21}$$

where

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \qquad \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

We denote by $\delta(t)$ the function which vanishes in W and is 1 in D, we denote by Ω the intersection of D with the circle $|z - z_0| < \rho$, and we denote by l the arc of L inside the circle $|z - z_0| < \rho$. Letting $\eta \to 0$, we obtain from (21) for $|t - z_0| < \rho$

$$(22) \quad (\mu + 1)\int_l \frac{y\,d\bar{z}}{z - t} + 4\iint_\Omega \psi_z\psi_{\bar{z}}\frac{dx\,dy}{y^2(z - t)} - \frac{16\pi}{t - \bar{t}}\psi_t^2\,\delta(t) + Q(t) = 0,$$

where $Q(t)$ is an analytic function in the circle $|t - z_0| < \rho$. Successive applications of Green's theorem to the integral over Ω in (22) show in turn that

$$(23) \quad (\mu + 1)\int_l \frac{y\,d\bar{z}}{z - t} + \delta(t)\left[-\frac{16\pi}{t - \bar{t}}\psi_t^2 + \frac{16\pi}{(t - \bar{t})^2}\psi\psi_t - \frac{4\pi}{(t - \bar{t})^3}\psi^2\right] + q(t) = 0,$$

where $q(t)$ is continuous for $|z - t| < \rho$.

In order to find the boundary values of ψ_t on l from (23), we introduce on l the function

$$w(z) = \int_{z_0}^{z} y\,d\bar{z}$$

defined by integration along l. The derivatives

$$(24) \qquad \dot{z} = \frac{\partial z}{\partial s}, \qquad \dot{w} = \frac{\partial w}{\partial s} = y\dot{\bar{z}}$$

exist almost everywhere on the monotonic arc l. We let $z_1 = x_1 + iy_1$ be a point of l where (24) holds, and we set $t_1 = z_1 + \xi e^{i\theta}$, $t_2 = z_1 - \xi e^{i\theta}$ for small $\xi > 0$, with $\theta = \pi/4$ if $x_1 > 0$ and $\theta = 3\pi/4$ if $x_1 < 0$. Thus t_1 lies in Ω, and t_2 lies in W. From (23) and an integration by parts we obtain

$$\begin{aligned}(25) \quad &\lim_{t_1 \to z_1} \frac{16\pi}{(1 + \mu)(t_1 - \bar{t}_1)}\psi_{t_1}^2 \\ &= \lim_{\xi \to 0}\left(\int_l \frac{y\,d\bar{z}}{z - t_1} - \int_l \frac{y\,d\bar{z}}{z - t_2}\right) \\ &= 2\pi i y_1 \dot{\bar{z}}_1^2 + \lim_{\xi \to 0} 2\int_l \left[\frac{w(z) - w(z_1)}{z - z_1} - y_1\dot{\bar{z}}_1^2\right]\frac{(t_1 - t_2)(z - z_1)^2\,dz}{(z - t_1)^2(z - t_2)^2}.\end{aligned}$$

One can check that the last limit appearing in (25) is zero by introducing a new coordinate system σ, τ with origin at z_1 and with the τ-axis inclined at an angle θ with the x-axis. Indeed, because of the monotonicity of l, the last integral (25) can be estimated easily [3] in terms of integrals with respect to σ over the intervals $(-\infty, -\epsilon^{\frac{1}{2}})$, $(-\epsilon^{\frac{1}{2}}, -\epsilon)$, $(-\epsilon, \epsilon)$, $(\epsilon, \epsilon^{\frac{1}{2}})$, and $(\epsilon^{\frac{1}{2}}, \infty)$.

Thus we see that almost everywhere along L

$$\lim \frac{2\psi_z}{y} = -i(1+\mu)^{\frac{1}{2}}\bar{z} \tag{26}$$

as z approaches L along a line inclined at 45° with the horizontal. The same calculation shows, with slight modifications, that ψ_z is bounded in the neighborhood of every point of L, except only for the intersection of L with the y-axis, near which L is not monotonic. However, the Phragmén-Lindelöf theorem can be used to show the boundedness of ψ_z even at that intersection [3]. The boundedness of the gradient of ψ together with the limit relation (26) almost everywhere along L constitutes a suitable generalized boundary condition from which we can study the free curve L without any further reference to the extremal problem (7).

4. The functional equation. We turn next to a proof that the free curve L is analytic. Our procedure [3] is the outgrowth of earlier work on free-surface flows in a gravity field [6].

The elliptic partial differential equation (2) has a fundamental solution S of the form

$$S(z,\bar{z};\zeta,\bar{\zeta}) = A(z,\bar{z};\zeta,\bar{\zeta}) \log (z-\zeta)(\bar{z}-\bar{\zeta}) + B(z,\bar{z};\zeta,\bar{\zeta}) \tag{27}$$

with a logarithmic pole at $z = \zeta$, where A and B are regular analytic functions of $z,\bar{z}$ and $\zeta,\bar{\zeta}$ and

$$A(z,\bar{z};z,\bar{z}) \equiv -\tfrac{1}{2}. \tag{28}$$

The function $A(z,\bar{z};\zeta,\bar{\zeta})$ is merely the Riemann function of the partial differential equation of hyperbolic type

$$\psi_{z\bar{z}} + \frac{1}{2(z-\bar{z})}\psi_z - \frac{1}{2(z-\bar{z})}\psi_{\bar{z}} = 0, \tag{29}$$

which is equivalent to (2) in the real domain. We have for A the explicit representation [1]

$$A(z,\bar{z};\zeta,\bar{\zeta}) = \frac{(z-\bar{\zeta})^{\frac{1}{2}}(\zeta-\bar{z})^{\frac{1}{2}}}{2(\bar{\zeta}-\zeta)} F\left[-\frac{1}{2},-\frac{1}{2},1,\frac{(z-\zeta)(\bar{z}-\bar{\zeta})}{(z-\bar{\zeta})(\bar{z}-\zeta)}\right] \tag{30}$$

in terms of the hypergeometric series F.

If Γ denotes a closed curve in D, we find by Green's theorem

$$\frac{2\,\delta(\zeta)\psi(\zeta,\bar{\zeta})}{\zeta-\bar{\zeta}} = \frac{1}{2\pi i}\int_\Gamma \left[\psi \frac{\partial}{\partial n} S(z,\bar{z};\zeta,\bar{\zeta}) - S(z,\bar{z};\zeta,\bar{\zeta})\frac{\partial \psi}{\partial n}\right]\frac{ds}{y}, \tag{31}$$

where $\delta(\zeta)$ now denotes the function which is 1 inside Γ and 0 outside Γ. By performing a suitable limit process on Γ, we can extend (31) to the case where one arc of Γ coincides with an arc l of L. On l we can use the formula (26) to determine $\partial\psi/\partial n$, by the Lebesgue convergence theorem.

Let us write $\zeta = \xi + i\eta$, $\zeta^* = \xi - i\eta$, and let us continue the identity (31) analytically into the domain of two complex variables ξ and η. Such an analytic continuation shows that

$$\frac{2\,\delta(\zeta)\psi(\zeta,\zeta^*)}{\zeta - \zeta^*} = \frac{1}{2\pi i}\int_\Gamma \left[\psi \frac{\partial}{\partial n} S(z,\bar{z};\zeta,\zeta^*) - S(z,\bar{z};\zeta,\zeta^*)\frac{\partial \psi}{\partial n}\right]\frac{ds}{y} \tag{32}$$

for independent values of ζ and ζ^*, provided ζ and $\bar{\zeta}^*$ lie either both inside or both outside Γ. This identity even holds in the limiting case where $\bar{\zeta}^*$ lies on l and ζ lies either in Γ or outside Γ. We shall deal only with this latter situation of $\bar{\zeta}^*$.

The determination of the branch of $\log (z - \zeta)(\bar{z} - \zeta^*)$ which must be used in (32) changes by $2\pi i$ on the arc of l between $\bar{\zeta}^*$ and ζ when ζ crosses l, but it does not change on the remainder of Γ. Thus by calculating the jump of each side of (32) as ζ crosses l and by taking into account the conditions (4) and (26) satisfied by ψ on l, we find that for ζ on l, ψ has continuous boundary values given by

$$\frac{2\psi(\zeta,\zeta^*)}{\zeta - \zeta^*} = (1+\mu)^{\frac{1}{2}}\int_{\bar{\zeta}^*}^{\zeta} A(z,\bar{z};\zeta,\zeta^*)\,ds. \tag{33}$$

Also, we can differentiate (32) with respect to ζ and again calculate the jump across l. Using the monotonicity of l, we can thus show that almost everywhere along L the boundary condition

$$2\frac{\partial}{\partial\zeta}\frac{\psi(\zeta,\zeta^*)}{\zeta - \zeta^*} = (1+\mu)^{\frac{1}{2}}A(\zeta,\bar{\zeta};\zeta,\zeta^*)\dot{\bar{\zeta}} + (1+\mu)^{\frac{1}{2}}\int_{\bar{\zeta}^*}^{\zeta}\frac{\partial}{\partial\zeta}A(z,\bar{z};\zeta,\zeta^*)\,ds \tag{34}$$

is fulfilled. The formulas (33), (34) merely represent the solution ψ of the Cauchy problem for (29) in terms of the Riemann function A, with initial data given along L.

Motivated by (34), we set up for any fixed ζ^*, with $\bar{\zeta}^*$ on L, the system of two integral equations

$$f(\zeta) = \frac{2}{(1+\mu)^{\frac{1}{2}}A(\zeta,g(\zeta);\zeta,\zeta^*)}\frac{\partial}{\partial\zeta}\frac{\psi(\zeta,\zeta^*)}{\zeta - \zeta^*} - \int_{\bar{\zeta}^*}^{\zeta}\frac{A_\zeta(z,g(z);\zeta,\zeta^*)}{A(\zeta,g(\zeta);\zeta,\zeta^*)}f(z)\,dz, \tag{35}$$

$$g(\zeta) = \zeta^* + \int_{\bar{\zeta}^*}^{\zeta} f(z)^2\,dz, \tag{36}$$

for the determination, in D and on L near $\bar{\zeta}^*$, of the two analytic functions $f(\zeta)$ and $g(\zeta)$. The method of successive approximations applies to the system (35), (36) and can be used to prove [3, 7] the existence of a unique solution f,g near $\bar{\zeta}^*$ which is analytic and is independent of the path from $\bar{\zeta}^*$ to ζ. All that is required for this existence proof is the known boundedness of the derivative ψ_ζ.

We notice by (34) that along L,

$$f(\zeta) = \dot{\bar{\zeta}}, \qquad g(\zeta) = \bar{\zeta}$$

is a solution of (35) and (36). As an integral, $g(\zeta)$ must be continuous in $D + L$ near $\bar{\zeta}^*$, and by the uniqueness of the solution of (35) and (36) we conclude that the analytic function $g(\zeta)$ in D has on the free boundary L the values

$$g(\zeta) = \bar{\zeta}. \tag{37}$$

Let $t = t(\zeta)$, $\zeta = \zeta(t)$ be a conformal mapping of D onto the upper half of the t-plane, and let

$$\Phi(t) = \zeta(t) + g(\zeta(t)),$$
$$\Psi(t) = \zeta(t) - g(\zeta(t)).$$

Then on the interval of the real axis corresponding to L

$$\Phi(t) = \text{real}, \qquad \Psi(t) = \text{imaginary}, \tag{38}$$

by (37), and hence Schwarz's reflection principle allows us to continue Φ and Ψ across the real axis into the lower half plane. In particular,

$$\zeta(t) = \frac{\Phi(t) + \Psi(t)}{2} \tag{39}$$

is analytic for real values of t corresponding to L, and hence the free boundary L is an analytic arc.

It is now easy [3] to conclude that ψ is analytic on L and that the constant-pressure condition (11) is satisfied identically there.

5. Remarks on uniqueness and construction. Let D_1 and D_2 be two different cavitational flow regions obtained by solving the extremal problem (7) with two values μ_1 and μ_2 of the cavitation parameter. Suppose that we can magnify D_1 by a factor $\alpha \geqq 1$ until its lower boundary curve lies just inside D_2. Then we can show by the generalized maximum principle for (2) that

$$\alpha^2\psi_1\left(\frac{x}{\alpha}, \frac{y}{\alpha}\right) < \psi_2(x,y) \tag{40}$$

in the intersection of D_2 and the magnified region D_1, since by our lemma on symmetrization the level curve

$$\alpha^2\psi_1\left(\frac{x}{\alpha}, \frac{y}{\alpha}\right) = \frac{h_0^2}{2} \tag{41}$$

lies above the horizontal line $y = h_0$. Hence at a point of tangency of the free boundaries of D_2 and the magnified image of D_1

$$\frac{1}{y}\frac{\partial}{\partial n}\left[\alpha^2\psi_1\left(\frac{x}{\alpha}, \frac{y}{\alpha}\right)\right] < \frac{1}{y}\frac{\partial\psi_2}{\partial n}. \tag{42}$$

Thus we obtain for μ_1 and μ_2 the useful inequality [5]

$$\mu_1 < \mu_2. \tag{43}$$

From (43) we can conclude that the extremal cavitating flow obtained from (7) for a given value of μ is unique [2, 3, 4, 5], since otherwise we would

arrive at the absurd inequality $\mu < \mu$. It is also possible to deduce from (43) that the cavity W depends continuously on μ and increases steadily as μ increases [2, 3]. In particular, for one special value of the cavitation parameter the free boundary L is an analytic arc joining the ends of the segment $y = h$, $-k \leqq x \leqq k$. This statement indicates that we can prescribe the point of separation of the free boundary by suitable choice of μ.

A case of special interest for the existence theorem on cavity flow completed above is that in which the curve C consists outside the strip $-k < x < k$ of two vertical segments $x = -k$ and $x = k$, $0 \leqq y \leqq h$. This is the case of a nose and tail consisting of circular disks. We take the point of separation of L from C to be $k + ih$ and we set $\bar{\zeta}^* = k + ih$ in (35) and (36). We check easily that for ζ on the segment $x = k$, $0 \leqq y \leqq h$, the expression

$$2\frac{\partial}{\partial\zeta}\frac{\psi(\zeta,\zeta^*)}{\zeta - \zeta^*} = \frac{\partial}{\partial\zeta}\int_{\bar{\zeta}^*}^{\zeta} A(z,\bar{z};\zeta,\zeta^*)\frac{\partial\psi}{\partial n}\frac{ds}{y} \tag{44}$$

is pure imaginary. Furthermore, for pure imaginary values of $g - \zeta^*$,

$$A(\zeta,g;\zeta,\zeta^*) = \text{real}, \tag{45}$$

$$A_\zeta(z,g;\zeta,\zeta^*) = \text{imaginary} \tag{46}$$

on that segment. Hence there is a solution f,g of (35) and (36) along the segment $x = k$, $0 \leqq y \leqq h$, with

$$g - \zeta^* = \text{imaginary}, \qquad f = \text{imaginary} \tag{47}$$

there. By the uniqueness of the solution of (35), (36) we deduce that the original analytic function $g(\zeta)$ satisfying (37) has on that segment a pure imaginary variation.

This information together with (38) would enable us to map the Φ-plane upon the Ψ-plane and thus determine L explicitly if we could make similar deductions along the x-axis. As it is, we see that on C and L the variation of Ψ is imaginary, while on C the variation of Φ is imaginary and on L the variation of Φ is real. Thus by studying the conformal mapping of the Φ-plane onto the Ψ-plane we can obtain an asymptotic development for Ψ as a function of Φ near the point of separation $\bar{\zeta}^*$. This development can be used to show that the local behavior of the free boundary at the point of separation from a vertical fixed boundary is the same for axially symmetric flow as for plane flow, since the functions g, Φ, and Ψ, defined by (37), have the same behavior in the plane case as that described above.

Finally, the equations (33), (35), (36), and (37) allow us to construct explicitly an example of axially symmetric cavitational flow with an arbitrary given analytic curve L as free boundary. Indeed, given an analytic curve L we can write down explicitly an expression for the analytic function $w = g(\zeta)$ with the property (37) on that curve. For points ζ and $\bar{\zeta}^*$ on L we set

$$\psi(\zeta,\zeta^*) = \frac{(\zeta - \zeta^*)(1+\mu)^{\frac{1}{2}}}{2}\int_{\zeta^*}^{\zeta} A(z,g(z);\zeta,\zeta^*)g'(z)^{\frac{1}{2}}\,dz, \tag{48}$$

a step which is motivated by (33), (35), and (36). We note that $g(\bar{\zeta}^*) = \zeta^*$, and we let $\zeta = G(w)$ be the inverse of $w = g(\zeta)$. From (48) we have

$$\psi(\zeta,w) = \frac{(\zeta - w)(1+\mu)^{\frac{1}{2}}}{2} \int_{G(w)}^{\zeta} A(z,g(z);\zeta,w)g'(z)^{\frac{1}{2}}\,dz, \tag{49}$$

replacing ζ^* by w. Formula (49) is an analytic identity for ζ and w on L, and hence by analytic continuation it should hold in some regions of the z-plane and the w-plane. In particular, we find [3] that

$$\psi(\zeta,\bar{\zeta}) = \frac{(\zeta - \bar{\zeta})(1+\mu)^{\frac{1}{2}}}{2} \int_{G(\bar{\zeta})}^{\zeta} A(z,g(z);\zeta,\bar{\zeta})g'(z)^{\frac{1}{2}}\,dz \tag{50}$$

is a real solution of (2) which represents the stream function of an actual axially symmetric cavitational flow with the given analytic curve L as its free boundary.

Bibliography

1. G. Darboux, *Leçons sur la théorie générale des surfaces*, Vol. 2, Paris, 1915.
2. P. R. Garabedian and D. C. Spencer, *Extremal methods in cavitational flow*, J. Rational Mech. Analysis vol. 1 (1952) pp. 359–409.
3. P. R. Garabedian, H. Lewy, and M. Schiffer, *Axially symmetric cavitational flow*, Ann. Math. vol. 56 (1952) pp. 560–602.
4. D. Gilbarg, *Uniqueness of axially symmetric flows with free boundaries*, J. Rational Mech. Analysis vol. 1 (1952) pp. 309–320.
5. M. Lavrentieff, *On certain properties of univalent functions and their application to wake theory* (in Russian), Rec. Math. (Mat. Sbornik) N.S. vol. 46 (1938) pp. 391–458.
6. H. Lewy, *A note on harmonic functions and a hydrodynamical application*, Proc. Amer. Math. Soc. vol. 3 (1952) pp. 111–113.
7. ———, *A theory of terminals and the reflection laws of partial differential equations*, Tech. Report No. 4, Stanford University, 1952.
8. L. E. Payne and A. Weinstein, *Capacity, virtual mass and generalized symmetrization*, Pacific J. Math. vol. 2 (1952) pp. 633–641.
9. G. Pólya and G. Szegö, *Isoperimetric inequalities of mathematical physics*, Princeton, 1951.
10. D. Riabouchinsky, *Sur un problème de variation*, C. R. vol. 185 (1927) pp. 840–841.

Stanford University,
Palo Alto, Calif.

REVIEW OF SIGNIFICANT OBSERVATIONS ON THE MACH REFLECTION OF SHOCK WAVES

BY

WALKER BLEAKNEY

The problem of the interaction of one shock front with another has received considerable attention in recent years and in the restricted form of the oblique reflection of a plane wave from a rigid wall has been the object of both theoretical and experimental attack. In spite of the careful investigations which have thus far been made many features of the phenomena associated with the problem are not understood. The situation has been summarized in several recent papers [3, 4, 16]. It is the purpose of this communication to discuss the salient observational features of Mach reflection which must be taken into account in any complete theory of the phenomenon. In writing this short account some known points are reviewed and some new ones added, and it will be assumed that the reader is familiar with the main ideas and notation given in recent discussions [3, 16].

Mathematical difficulties have excluded, as yet, a solution of the hydrodynamical problem of a plane shock of finite strength in an inviscid but compressible fluid striking symmetrically a wedge of finite angle. In the absence of a complete solution a local three-shock theory was suggested by von Neumann [2] to apply in the neighborhood of the intersection of the incident, reflected, and Mach shock fronts. Essentially the same ideas give a solution for regular reflection (two-shock theory) which corresponds with the experimental observations. An unexpected difficulty appeared, however, when Smith's results [6] indicated that the three-shock configurations were found in angular domains where no solutions of the three-shock equations exist.

Stimulated by this discrepancy, a number of asymptotic solutions of the hydrodynamical problem valid under limiting conditions have been found. The work of Bargmann [9], Lighthill [19], and Ting and Ludloff [20] covers the case of all shock strengths at nearly glancing incidence and Keller and Blank [24] the case of vanishing shock strength at all angles of incidence, while Lighthill [26] has also treated the case of nearly head-on reflection. In so far as these solutions can be compared with experiment no disagreement has been found. Where Mach reflection appears in these asymptotic solutions, the reflected wave is vanishingly weak and there is no contradiction with the local three-shock theory. It is also worth noting that the assumption of a perfect fluid in the asymptotic solutions seems to be quite justified.

The difficulty in explaining Mach reflection at finite angles has led to many speculations regarding the validity of the experimental measurements and the significance of certain observed features. So far no departure from the pseudo-stationary state has been detected (*i.e.*, the whole phenomenon may be described in terms of x_1/t and x_2/t, for example, where x is a coordinate and

t the time). To give an order of magnitude when t is changed by a factor of 2, the change in x/t is certainly less than 1 per cent. Regarding the angles made by the four discontinuities meeting at the triple point, there are rather large uncertainties because of the curvature of some of the shocks, but the incident shock and the wall are plane, and the path of the triple point is a straight line, so that the angles between these three can be measured with

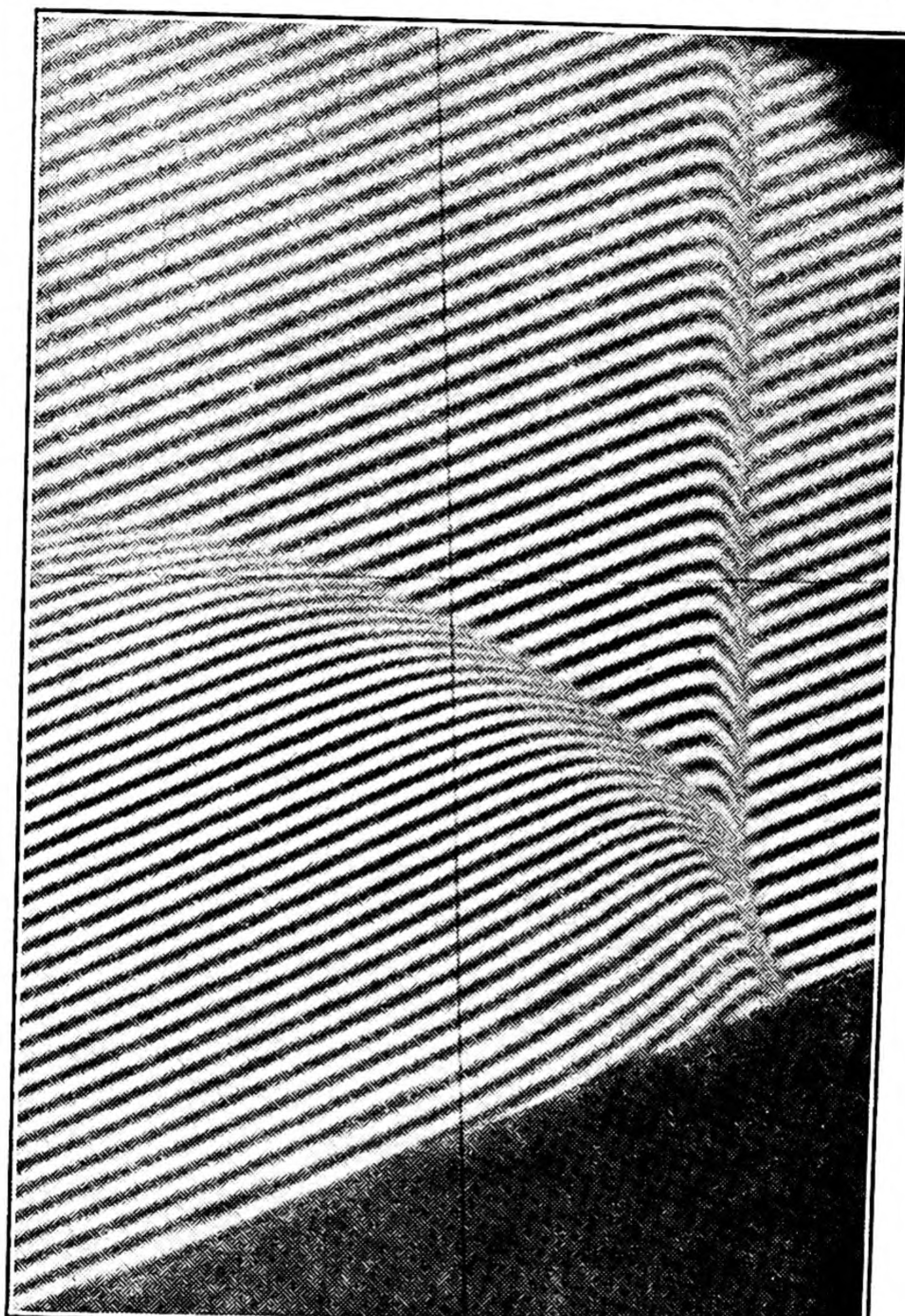


FIG. 1. Mach reflection in CO_2. The fringes shift discontinuously at the shock fronts and then continuously through the relaxation zone. Incident pressure ratio $\xi = 0.55$, angle of incidence $\alpha = 68°$.

good precision, and these alone suffice to show that the local three-shock theory does not fit.

Most of the experimental observations have been made in air, which is assumed to obey the perfect-gas law. An attempt was made to see whether the nature of the gas had any marked influence on Mach reflection by repeating the experiment in argon and carbon dioxide. The first result had all the appearances observed in air, while the second showed large relaxation effects (Fig. 1), but in neither case could the results be fitted to the simple theory.

The treatment of the problem neglects the viscosity and heat conductivity of the gas, and these points obviously require examination. Unfortunately it is difficult experimentally to obtain gases with sufficient variation in these respects to throw any light on the question. The strongest arguments in justification of the assumptions of a perfect fluid are first the pseudo-stationary property and second the success of the two-shock theory of regular reflection without the necessity of introducing transport phenomena.

Interferometer measurements of the density [16] have failed to reveal any angular variations such as Prandtl-Meyer expansions which might alter the

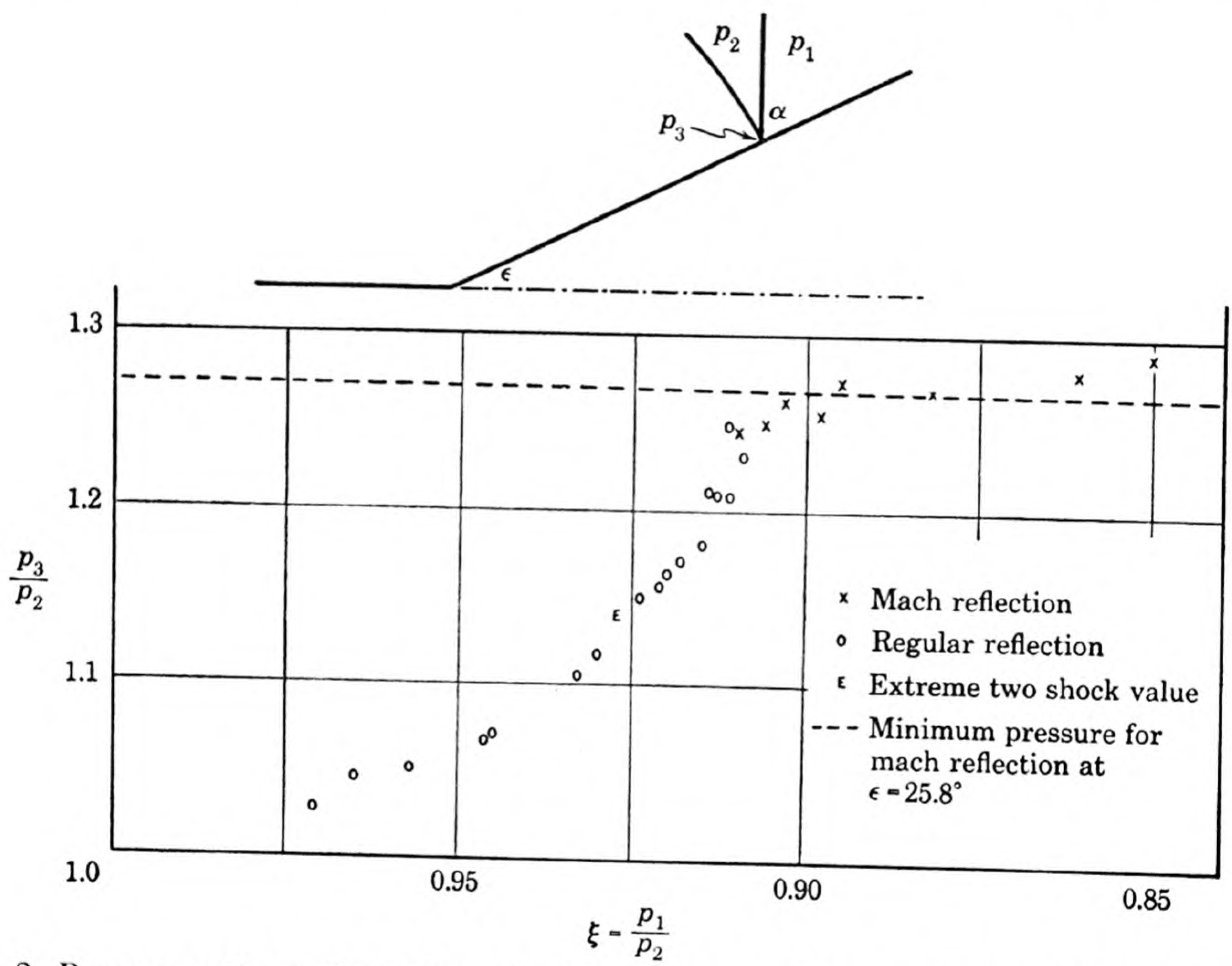


FIG. 2. Pressure ratio at the wall in the transition zone from regular to Mach reflection.

nature of the singularity at the triple point. Another anomaly first observed by Smith [6] is the apparent persistence of regular reflection somewhat beyond the extreme angle where the two-shock equations cease to apply. This has been confirmed by White [23] in a different way by measuring the pressure just behind the intersection at the wall and plotting this result as a function of pressure ratio ξ across the incident shock (Fig. 2). The horizontal dotted line represents the steady Mach solution, *i.e.*, the pressure behind a shock perpendicular to the wall traveling with the speed of the intersection. The conclusion is that within the accuracy of the experiments Mach reflection does not begin until the pressure is sufficiently great to support the Mach wave, and this occurs at a point beyond the pressure ratio corresponding to the extreme angle.

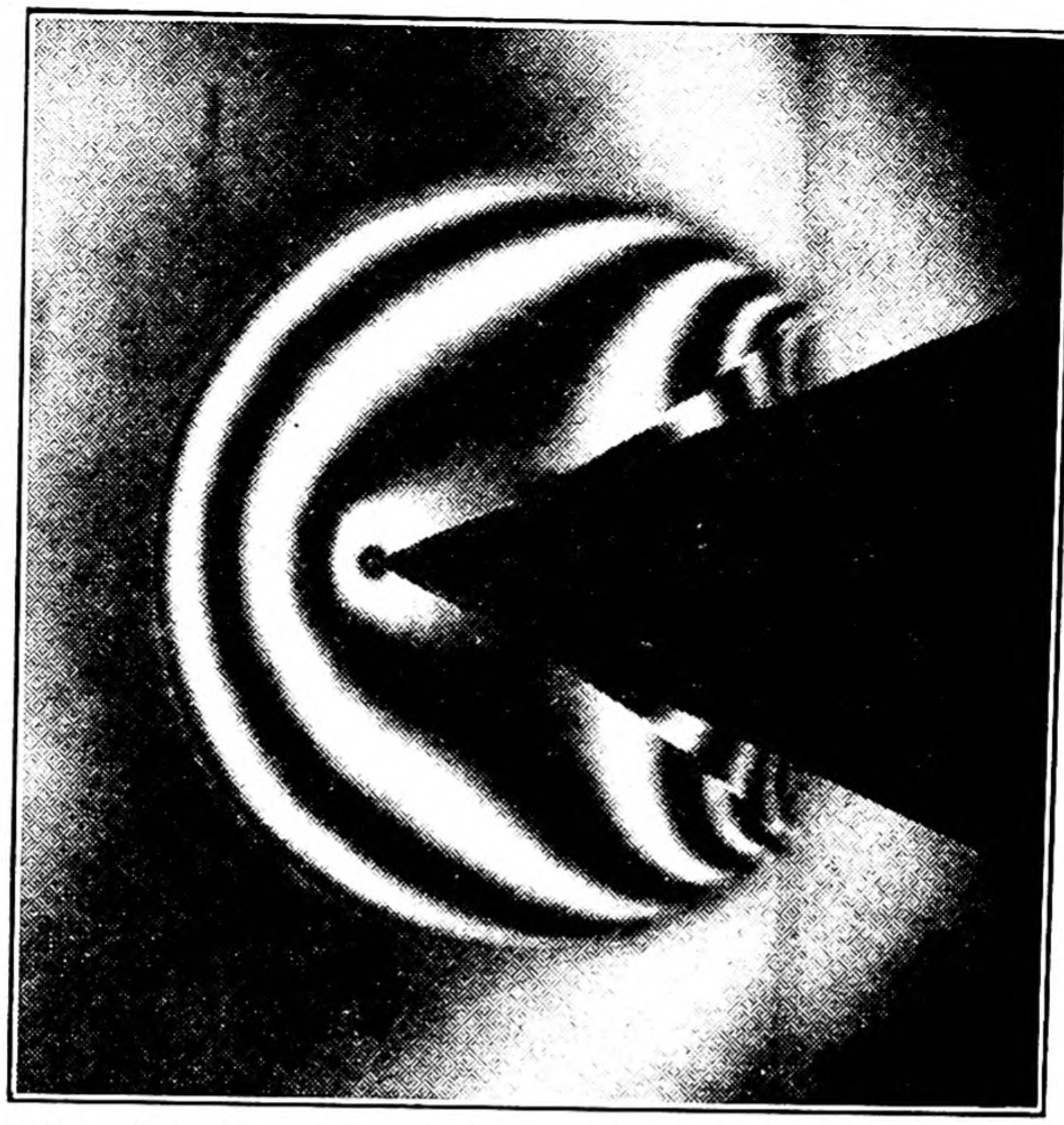


FIG. 3. Contour-fringe interferogram of the complete field of disturbance in the reflection from a wedge.

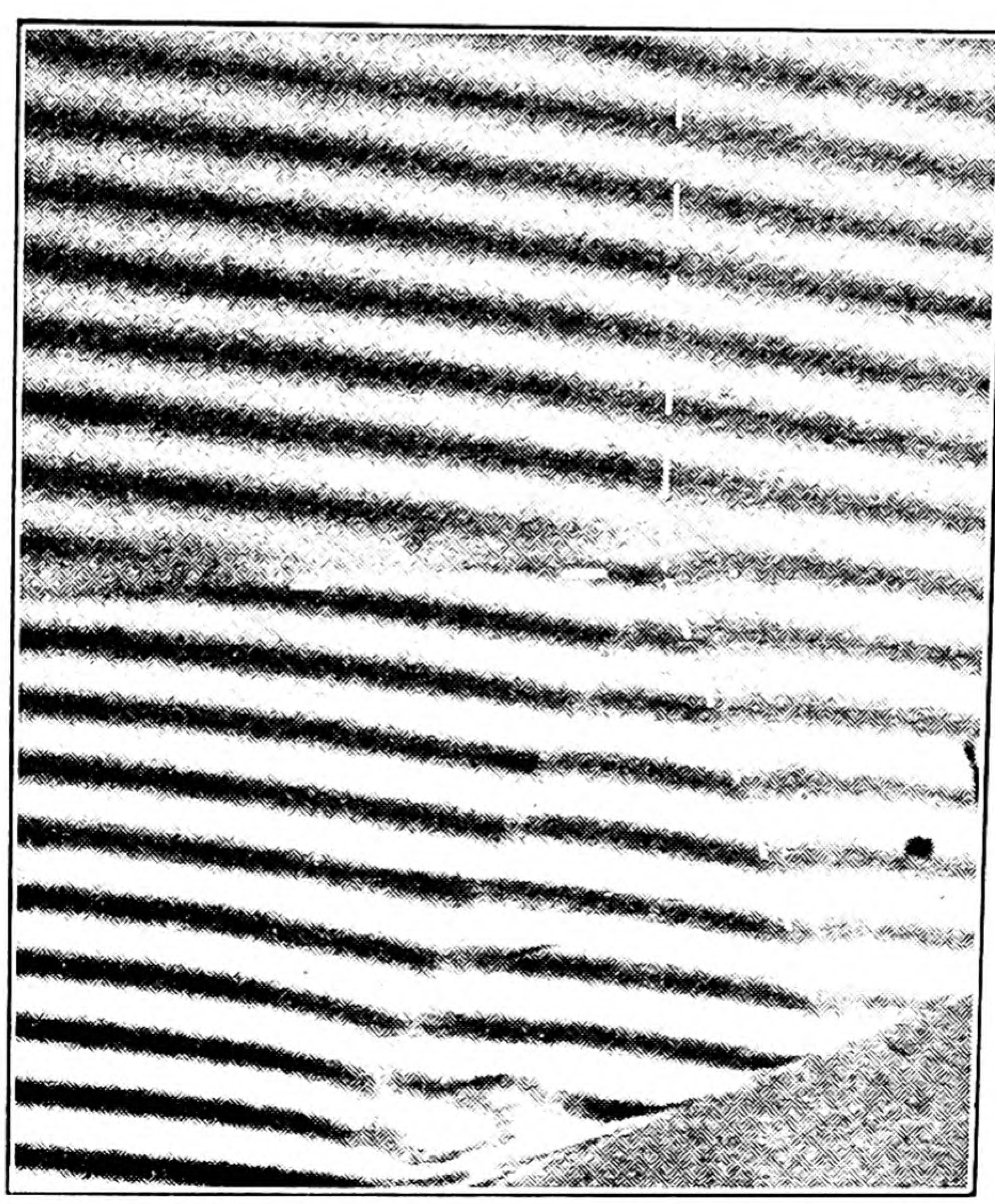


FIG. 4. The curling slip stream as it meets the wall.

No clue to the puzzle in the neighborhood of the shock intersection having been found, it is appropriate to look at other regions of the disturbance. A question that arises concerns the nature of the corner where the reflection first begins. In some of the early work this corner was represented by the junction of an inclined plate with the wall of the shock tube, and investigation showed that there was an interaction here with the boundary layer at the later times. To eliminate this disturbance as far as possible, a symmetrical wedge was used in all the later work, as shown in Fig. 3, but no detectable change in the reflection process itself was observed.

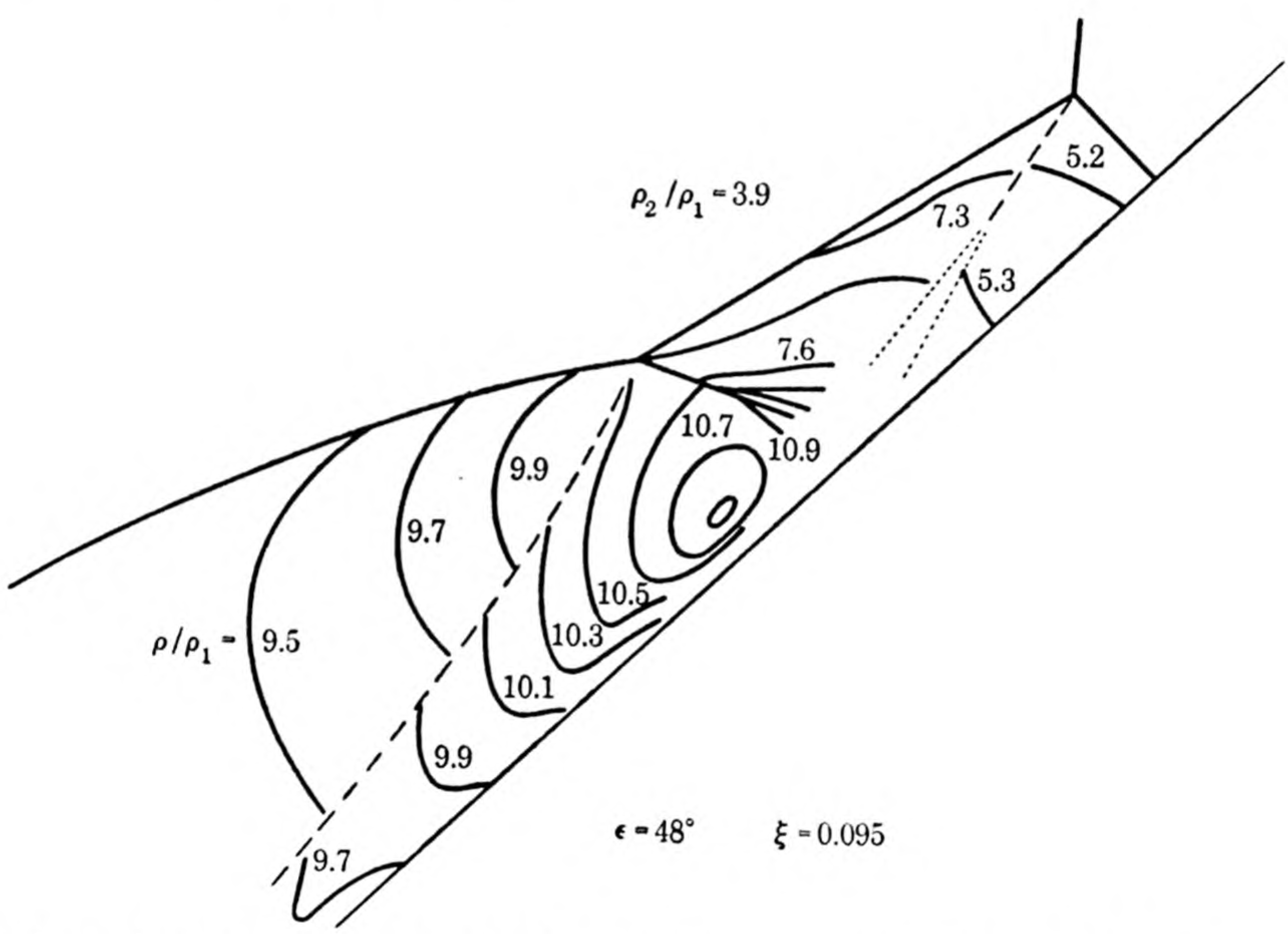


FIG. 5. Secondary shock in Mach reflection of very strong waves. The slip stream is turbulent near the wall. Note the second slip stream. Density contours are numbered. The densities in front of and behind the incident shock are denoted by ρ_1 and ρ_2, respectively. Angle of incidence $\alpha = 42°$.

The boundary conditions on the slip stream as it meets the wall are such that either it is tangent to the wall or its intensity vanishes there. Some adjustment of the stream was therefore expected, but the manner in which this takes place was not foreseen. As shown in the interferogram in Fig. 4, the end of the slip stream near the wall curls up on itself in a peculiar manner, the effect showing up best for strong shocks of pressure ratios in the neighborhood of 10. In this range the reflected and Mach shocks are quite straight near the triple point, and the three-shock theory fits pretty well. As the experiments are carried to stronger shocks or smaller angles of incidence, a point of inflection, or kink, begins to appear in the reflected wave [23] and a compression forms along a front extending roughly from the kink to the point at which the slip stream meets the wall. In the extreme case the point of inflection seems

to develop into a much higher order of contact, the compression becomes a shock, and a second slip stream is visible, as shown in Fig. 5.

For some time the fact has been emphasized that the three-shock theory failed for weak shocks rather than strong. Perhaps a more significant way of stating the same fact is that the theory fails for curved shocks. Some effects of curvature and its derivatives have been worked out by Taub [31], and these may throw some light on the local theory.

BIBLIOGRAPHY

1. E. Mach, Akad. Wiss. Wien vol. 77 (1878) p. 819.

2. J. von Neumann, *Oblique reflection of shocks*, E.R.P. No. 12, Buord. U.S. Navy Department (October, 1943).

3. W. Bleakney and A. H. Taub, *Interaction of shock waves*, Rev. Mod. Phys. vol. 21 (1949) p. 584.

4. H. Polachek and R. J. Seeger, *On shock-wave phenomena; interaction of shock waves in gases*, Proc. Symposia Appl. Math. Amer. Math. Soc. vol. 1 (1949).

5. W. Bleakney, *An experimental study of the interaction of shock waves*, Naval Ordnance Laboratory Rep. 1135 (June, 1949).

6. L. G. Smith, *Photographic investigations of the reflection of plane shocks in air*, O.S.R.D. No. 6271 (1945).

7. H. Polachek and R. J. Seeger, *Regular reflection of shocks in ideal gases*, E.R.P. No. 13, Buord. U.S. Navy Department (1944).

8. W. Bleakney, C. H. Fletcher, and D. K. Weimer, *The density field in Mach reflection of shock waves*, Phys. Rev. vol. 76 (1949) p. 323.

9. V. Bargmann, *On nearly glancing reflection of shocks*, O.S.R.D. No. 5171 (1945). Condensed in [16].

10. F. B. Harrison and W. Bleakney, *Remeasurement of reflection angles in regular and Mach reflection of shock waves*, Tech. Rep. II-0, Princeton Univ. (March, 1947).

11. A. H. Taub, *Refraction of plane shock waves*, Phys. Rev. vol. 72 (1947) p. 51.

12. ———, *Some numerical results on refraction of plane shocks*, Tech. Rep. II-1, Princeton Univ. (March, 1947).

13. C. H. Fletcher, D. K. Weimer, and W. Bleakney, *The pressure behind a shock wave diffracted through a small angle*, Phys. Rev. vol. 78 (1950) p. 634.

14. C. H. Fletcher, *The Mach reflection of weak shock waves*, Tech. Rep. II-4, Princeton Univ. (June, 1950). Material condensed in [16].

15. D. R. White, *An experimental survey of the Mach reflection of shock waves*, Tech. Rep. II-10, Princeton Univ. (August, 1951). Material condensed in [23].

16. C. H. Fletcher, A. H. Taub, and W. Bleakney, *The Mach reflection of shock waves at nearly glancing incidence*, Rev. Mod. Phys. vol. 23 (1951) p. 271.

17. D. R. White, *Reflection of strong shocks at nearly glancing incidence*, J. Aero. Sci. vol. 18 (1951) p. 633.

18. G. Rawling and J. Riley, *Glancing reflection of shock*, U.S. Naval Ordnance Rep. 2165 (February, 1952).

19. M. J. Lighthill, *The diffraction of blast I*, Proc. Roy. Soc. London Ser. A vol. 198 (1949) p. 454.

20. L. Ting and N. F. Ludloff, *Aerodynamics of blast*, J. Aero. Sci. vol. 18 (1951) p. 143.

21. H. Polachek and R. J. Seeger, *On shock-wave phenomena; refraction of shock waves at a gaseous interface*, Phys. Rev. vol. 84 (1951) p. 922.

22. R. Courant and K. O. Friedrichs, *Supersonic flow and shock waves*, Interscience Publishers, Inc., New York, 1948.

23. D. R. White, *An experimental survey of the Mach reflection of shock waves*, Second Mid-West Conference on Fluid Dynamics (March, 1952).

24. J. B. Keller and A. Blank, *Diffraction and reflection of pulses by wedges and corners*, Comm. Pure Appl. Math. vol. 4 No. 1 (1951) p. 75.

25. O. Kay, *Diffraction of an arbitrary pulse by a wedge*, New York Univ. Research Rep. EM-43 (April, 1952).

26. M. J. Lighthill, *Diffraction of blast II*, Proc. Roy. Soc. London Ser. A vol. 200 (1950) p. 554.

27. D. M. Jones, P. M. E. Martin, and C. K. Thornhill, *A note on the pseudo-stationary flow behind a strong shock diffracted at a corner*, Proc. Roy. Soc. London Ser. A vol. 209 (1951) 238.

28. H. S. Tan, *Strength of reflected shock in Mach reflection*, J. Aero. Sci. vol. 18 (1951) p. 768.

29. G. H. Lean, *Report on further experiments on the reflection of inclined shock waves*, National Physical Laboratory, London, 1946.

30. V. Bargmann and D. Montgomery, *Prandtl-Meyer zones in Mach reflection*, O.S.R.D. No. 5011*b* (April, 1945).

31. A. H. Taub, *Curved shocks in pseudo-stationary flows* (to be published).

32. A. Sommerfeld, *Theoretisches über die Bengung der Röntgenstrahlen*, Zeit. Mach. Phys. vol. 46 (1901) p. 11.

33. F. G. Friedlander, *The diffraction of sound pulses*, I, *Diffraction by a semi-infinite plane*, II, *Diffraction by an infinite wedge*, III, *Note on an integral occurring in I*, IV, *On a paradox in the theory of reflection*, Proc. Roy. Soc. London Ser. A vol. 186 (1946) pp. 322, 344, 352, 356.

34. F. Wecken, *Problèmes pseudostationnaires*, Rep. Laboratoire d'Etudes Ballistiques de St. Louis, France (1947).

PRINCETON UNIVERSITY,
PRINCETON, N.J.

ON A NONLINEAR DIFFERENTIAL EQUATION IN HYDRAULICS

BY

N. W. McLACHLAN

1. Introduction. The equation

$$\ddot{u} + 2\kappa|u|\dot{u} + au = d \qquad (a,\kappa > 0, d \geqq 0) \tag{1}$$

occurs in connection with a hydroelectric power system. Referring to Fig. 1, water is supplied from a reservoir via a tunnel and pipe line to turbines which drive electrical generators. The rate of flow of water to the turbines is controlled by valves at the lower end of the pipe line, which may be opened or closed as desired. If one or more valves are closed, in order to avoid the occurrence of a large force in the system due to change in momentum of the surplus water, a surge tank is situated at the lower end of the tunnel. On closing one or more valves the water level in the tank rises above its value for steady flow, and this restricts the force to a safe maximum. When one or more

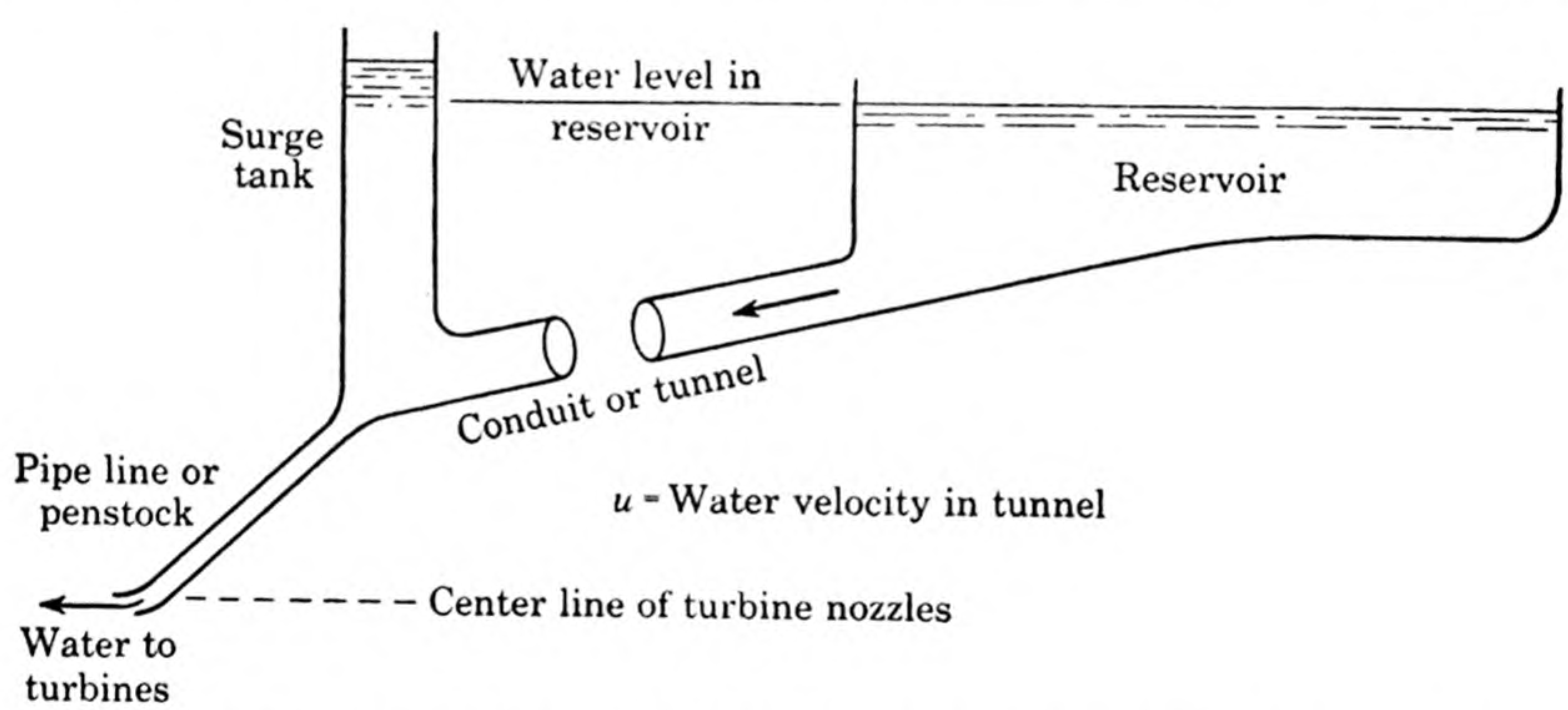


FIG. 1. Schematic diagram for hydroelectric system having surge tank.

valves are opened, water flows from the tank, thereby augmenting the supply from the reservoir. In either case, u, the velocity of the water in the tunnel, ultimately attains a substantially steady value $u = d/a$. Any change in the rate of flow of water to the turbines is accompanied by a surge in the tank, tunnel, and reservoir, which combination may be regarded as a form of enormous U tube having a nonuniform cross section. The surface area of the water in the reservoir far exceeds that in the tank, and so the level in the former is substantially constant during a surge. Owing to friction between the water and the internal surface of the U, loss occurs which varies as u^2. Since the friction opposes the motion of the water, irrespective of its direction, the damping *coefficient* is always positive and is represented by $2\kappa|u|$. The term au arises in virtue of the gravitational restoring force when the water level in the tank differs from its equilibrium value. For a tunnel of given

dimensions, $a \propto 1/A_s$, where A_s (assumed constant) is the cross-sectional area of the surge tank. d is proportional to Q, the rate of water flow (cubic feet per second) to the turbines. Thus $d = 0$ signifies that all the valves to the turbines are closed.

2. Alternative form of equation (1). Write $u = x + c$, where $c = d/a$, and (1) becomes

(2) $$\ddot{x} + 2\kappa|x + c|\dot{x} + ax = 0.$$

If the supply of water to the turbines is discontinued, $d = 0$ so $c = 0$, and (2) becomes

(3) $$\ddot{x} + 2\kappa|x|\dot{x} + ax = 0.$$

It may be shown for initial conditions $x(0) = x_0$, $\dot{x}(0) = x_1$, that x and $\dot{x} \to 0$ as $t \to +\infty$.

3. Singular points of (1). Putting $d = 0$ and writing $v = du/dt$, $v'_u = dv/du$, we obtain

(4) $$vv'_u + 2\kappa|u|v + au = 0,$$

so

(5) $$v'_u = \frac{-(au + 2\kappa|u|v)}{v}.$$

When $u = v = 0$, the right-hand side of (5) has the indeterminate form 0/0, so the origin is a singular point, and by Poincaré's criteria [1] it is either a center or a spiral point. Presence of the higher-order term $2\kappa|u|v$ does not permit discrimination between these two. However, in virtue of the last sentence in Sec. 2, the origin is a stable spiral point. Consequently, if $d = 0$, the water executes a damped oscillation whatever the value of $\kappa > 0$. This conclusion may be grasped more readily if it is realized that the damping coefficient $2\kappa|u| \to 0$ as $u \to 0$. Moreover, the motion is ultimately quasi-periodic with approximate period $2\pi/a^{\frac{1}{2}}$.

When $d > 0$, we use (2) in the form

(6) $$\frac{dv}{dx} = v'_x = \frac{-(ax + 2\kappa|x + c|v)}{v}.$$

Since $x \to 0$ as $t \to +\infty$, it follows that, at some $t = t_1 > 0$, $c > |x|$; so in $t_1 < t < \infty$, $|x + c|$ may be replaced by $x + c$.[1] Then Poincaré's criteria show that if $\kappa^2c^2 < a$, the point $x = v = 0$, that is, $u = c$, $v = 0$, is a stable spiral point, while if $\kappa^2c^2 \geqq a$, it is a stable node. In the first case the motion is a damped oscillation, whereas in the second it is damped but nonoscillatory (aperiodic). The general forms of the (v,y)-curves (obtained by the method of isoclines) for the three cases are depicted in Figs. 2, 3, and 4. The numbers on the thin curves indicate the slope of the (v,y)-curve at the crossing points.

[1] With certain initial conditions $c > |x|$ for $t_1 = 0$; see Sec. 4.

In connection with Fig. 2 it is of interest to remark that if the initial conditions are $|u_0| > 0$, $|v_0| = a/2\kappa$, then over the range $(|u_0|,0)$, we have $|\dot{u}| = |v| = a/2\kappa$, and so $|v|$ is constant, and the relationship between the water velocity u and time is linear. When u changes sign, the (u,t)-relationship ceases to be linear.

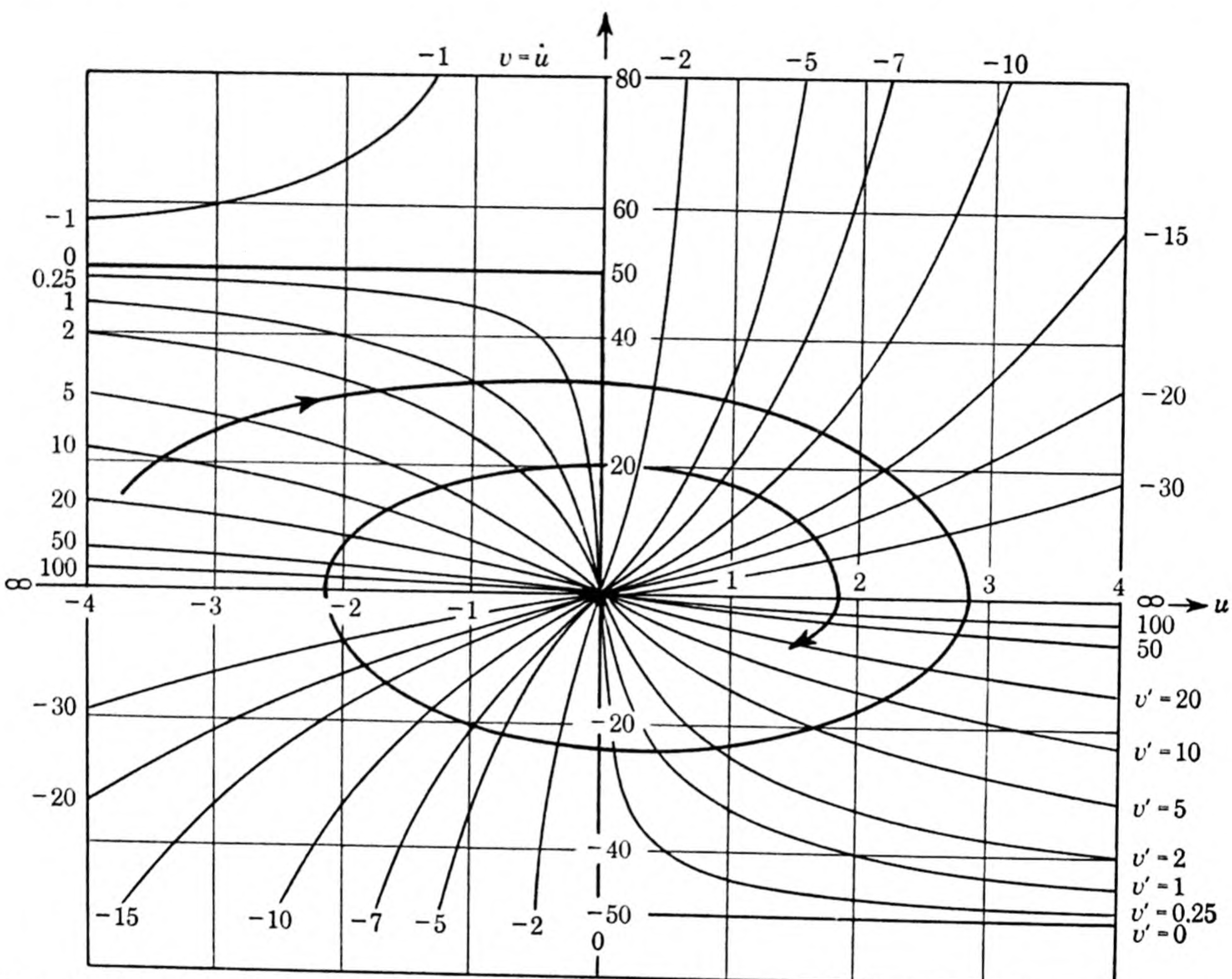


FIG. 2. Integral curve $(u,\dot{u})$ for $\ddot{u} + 2|u|\dot{u} + 100u = 0$. Oscillatory case with $d = 0$. The origin is a stable spiral point.

4. Solution of (2). If κ is such that the damping is not too large, we may employ the method of slowly varying amplitude and phase [2]. Then in the first approximation, the form of the solution is

$$u = A(t) \sin [\omega t + \varphi(t)], \tag{7}$$

$A(t)$ being the amplitude and $\varphi(t)$ the "phase" angle at any time $t \geqq 0$, that is, A and φ are functions of t whose rate of variation is small in comparison with ω. Writing $a = \omega^2$, $\omega t = \psi$, we have

$$\dot{A} = \frac{-\kappa A}{\pi} \int_0^{2\pi} |A \sin \psi + c| \cos^2 \psi \, d\psi. \tag{8}$$

Two cases may arise, namely, (1) $c = d/a \geqq |A|$ for all $t \geqq 0$; (2) $c = d/a < |A|$ for all $t \geqq 0$, or for $0 \leqq t < t_1$.

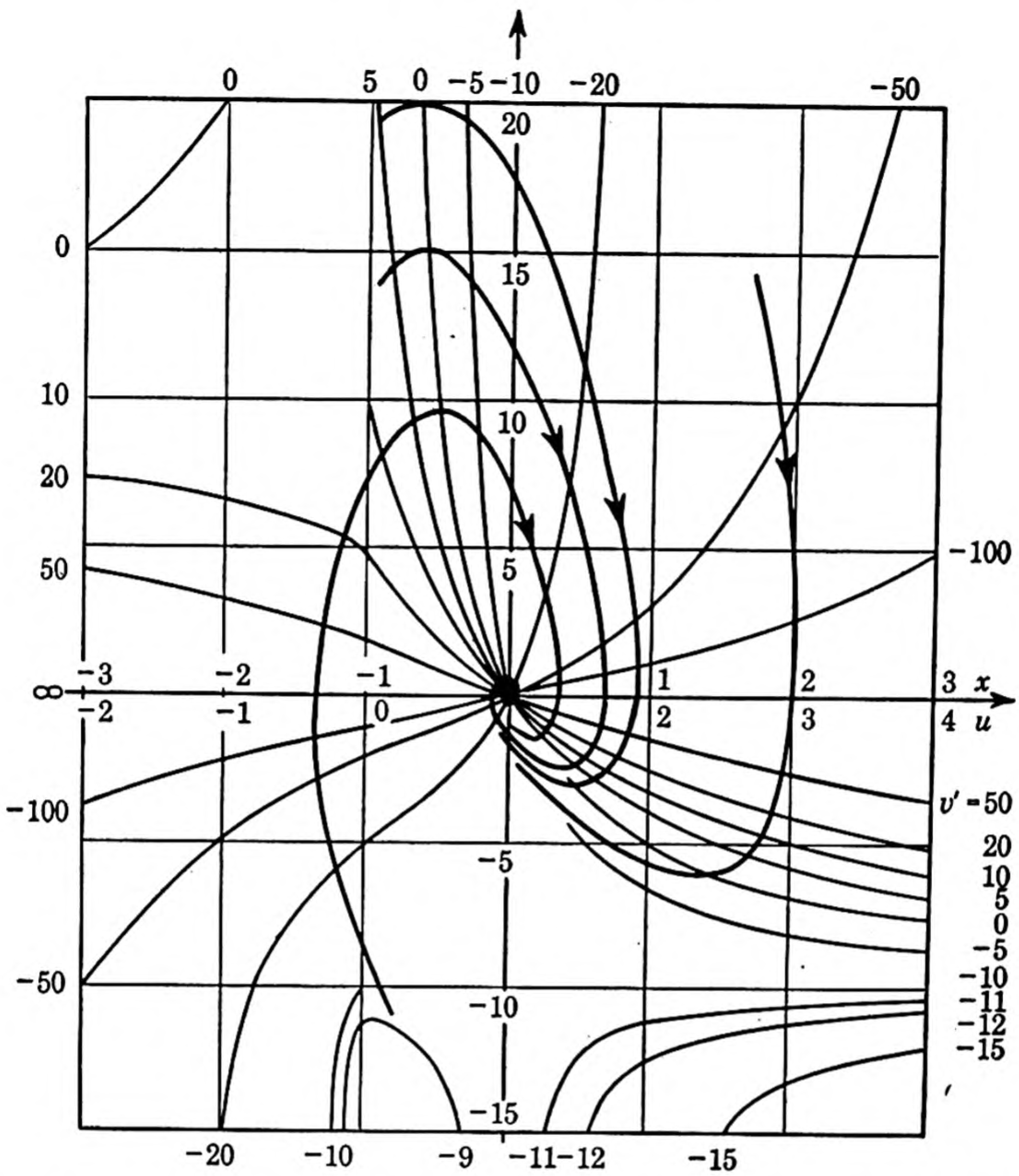


FIG. 3. Integral curves for $\ddot{u} + 10|u|\dot{u} + 100u = 100$, that is, $\ddot{x} + 10|x + 1|\dot{x} + 100x = 0$, with $u = x + 1$. Oscillatory case with $d = 100$ and $\kappa^2 c^2 < a$. $u = 1$, $v = 0$, is a stable spiral point.

Case 1. Here $|A \sin \psi + c|$ may be replaced by $(A \sin \psi + c)$, and since the integral corresponding to $A \sin \psi$ in (8) is zero, we obtain

$$\dot{A} = \frac{-\kappa c A}{\pi} \int_0^{2\pi} \cos^2 \psi \, d\psi = -\kappa c A, \tag{9}$$

and so

$$A(t) = A_0 e^{-\kappa c t}, \tag{10}$$

where A_0 is the amplitude at $t = 0$. Also we have

$$\dot{\varphi} = \frac{\kappa}{\pi} \int_0^{2\pi} (A \sin \psi + c) \cos \psi \sin \psi \, d\psi = 0, \tag{11}$$

so that

$$\varphi = \varphi_0, \text{ an } \textit{arbitrary} \text{ phase angle.} \tag{12}$$

Writing $\varphi_0 = (\pi/2) - \phi_0$, and substituting from (10), (12) into (7), the formal solution of (2) is

$$x = A_0 e^{-\kappa c t} \sin (\omega t + \varphi_0) = A_0 e^{-\kappa c t} \cos (\omega t - \phi_0). \tag{13}$$

Hence, finally, the velocity of the water in the tunnel in the first approximation is

$$u = x + c = A_0 e^{-\kappa c t} \cos (\omega t - \phi_0) + \frac{d}{a}, \tag{14}$$

A_0 and ϕ_0 being determinable from the initial conditions, *i.e.*, those which obtain at the beginning of a surge after d has been altered.

Initial conditions. Suppose that the steady rate of flow of water to the turbines is reduced from Q_1 to Q_2 by closing a valve.[2] Then as $t \to -0$, $u = c = d_1/a$, but for $t > 0$, $c = d_2/a$, and we have to determine the corresponding value of u. The initial conditions are $u(0) = d_1/a$, $\dot{u}(0) = 0$, with $c = d_2/a$ in (14). Thus

$$A_0 e^{-\kappa d_2 t/a} \cos (\omega t - \phi_0) \Big|_{t=0} + \frac{d_2}{a} = \frac{d_1}{a}, \tag{15}$$

so that

$$A_0 \cos \phi_0 = \frac{d_1 - d_2}{a}. \tag{16}$$

Differentiating (14) and putting $t = 0$ gives

$$\tan \phi_0 = \frac{\kappa d_2}{\omega a} \quad \text{and} \quad \cos \phi_0 = \left[1 + \left(\frac{\kappa d_2}{\omega a}\right)^2\right]^{-\frac{1}{2}}. \tag{17}$$

Substituting for $\cos \phi_0$ into (16) yields

$$A_0 = \frac{d_1 - d_2}{a}\left[1 + \left(\frac{\kappa d_2}{\omega a}\right)^2\right]^{\frac{1}{2}}. \tag{18}$$

Hence from (14) and (18) the velocity of the water in the tunnel during the surge is

$$u = \frac{d_1 - d_2}{a}\left[1 + \left(\frac{\kappa d_2}{\omega a}\right)^2\right]^{\frac{1}{2}} e^{-\kappa d_2 t/a} \cos (\omega t - \phi_0) + \frac{d_2}{a}, \tag{19}$$

where $\phi_0 = \tan^{-1} (\kappa d_2/\omega a)$. From (19) it is seen that as $t \to +\infty$, $u \to d_2/a$, the steady velocity after subsidence of the surge. By hypothesis $c = d_2/a \geqq A_0$; so we must have $d_1/\{1 + [1 + (\kappa d_2/\omega a)^2]^{\frac{1}{2}}\} < d_2 < d_1$.

Suppose that the steady rate of water flow to the turbines is *increased* from Q_1 to Q_2. Then $d_1 < d_2$, and if $d_2/a \geqq A_0$, (19) is valid and we have

$$u = -\frac{d_2 - d_1}{a}\left[1 + \left(\frac{\kappa d_2}{\omega a}\right)^2\right]^{\frac{1}{2}} e^{-\kappa d_2 t/a} \cos (\omega t - \phi_0) + \frac{d_2}{a}. \tag{20}$$

As before, when $t \to +\infty$, $u \to d_2/a$, the final unidirectional velocity of the water in the tunnel.

[2] The time taken for the change from Q_1 to Q_2 is assumed small compared with $2\pi/\omega$ [3].

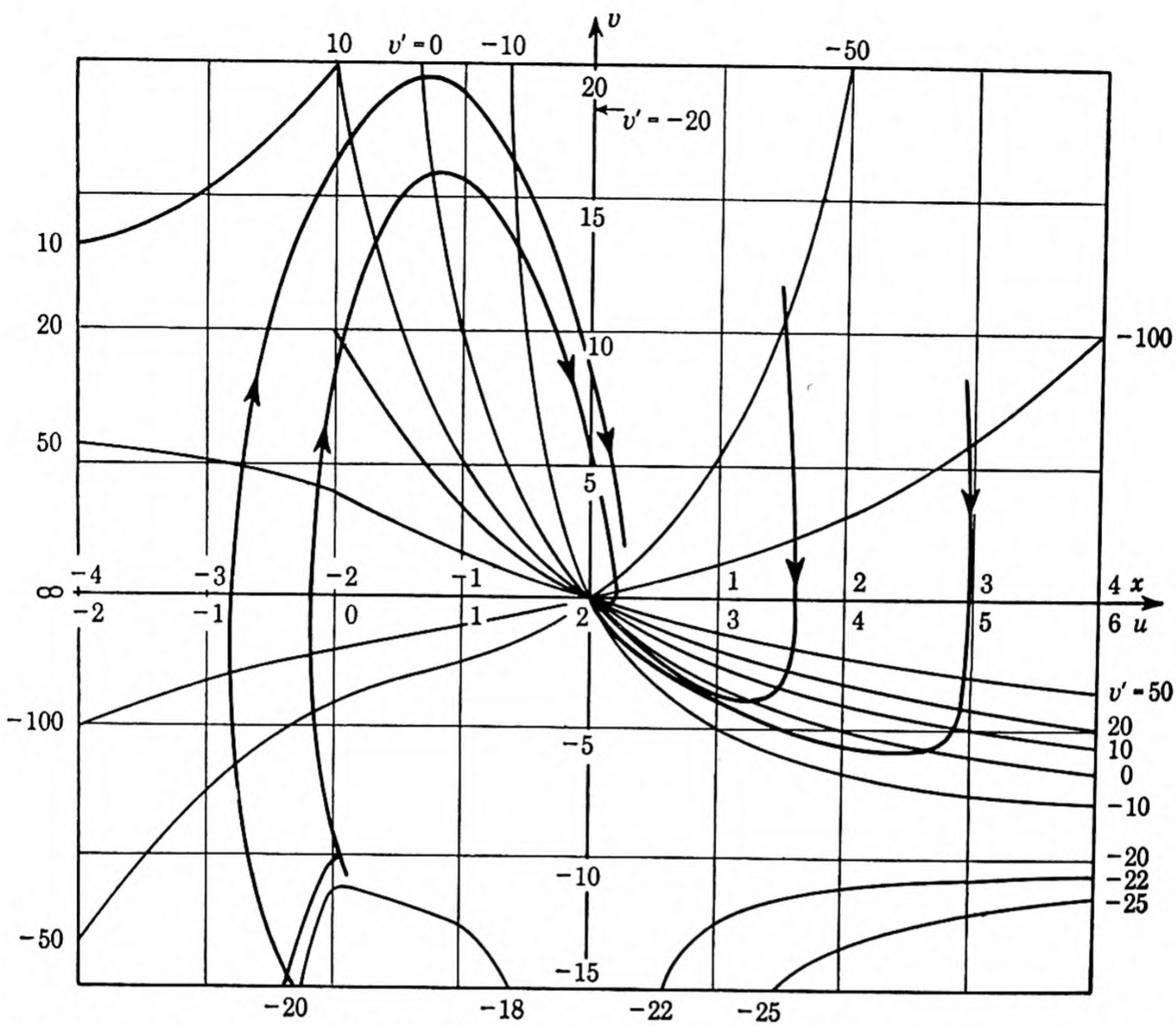


FIG. 4. Integral curves for $\ddot{u} + 10|u|\dot{u} + 100u = 200$, that is, $\ddot{x} + 10|x+2|\dot{x} + 100x = 0$. Nonoscillatory case with $d = 200$ and $\kappa^2c^2 = a$. $u = 2$, $v = 0$ is a stable node.

5. Case 2. $|A| > c = d/a$, $0 \leqq t < t_1$. In this instance evaluation of the integral in (8) yields

$$\dot{A} = -\frac{\kappa A}{\pi}\left[\left(\frac{2(A^2 - c^2)^{\frac{1}{2}}(2A^2 + c^2)}{3A^2}\right) + 2c\sin^{-1}\frac{c}{A}\right], \tag{21}$$

and as in Sec. 4 we assume that the damping is not too large. If $c/|A| \leqq 0.4$, as a first approximation (21) gives

$$\dot{A} = -\left(\frac{4\kappa}{3\pi}\right)A^2, \tag{22}$$

so that

$$A^{-1} = \frac{4\kappa t}{3\pi} + B,\ \text{a const.} \tag{23}$$

If $A = A_0$ when $t = 0$, $B = 1/A_0$ and

$$A(t) = \frac{A_0}{1 + (4A_0\kappa t/3\pi)}. \tag{24}$$

We find also that $\varphi = \varphi_0$, an arbitrary phase angle. Hence by (24) and (7), with $\varphi_0 = (\pi/2) - \phi_0$, we get

$$u = x + c = \frac{A_0}{1 + \gamma t} \cos (\omega t - \phi_0) + c, \tag{25}$$

where $\gamma = 4A_0\kappa/3\pi$, and provided that $A_0/(1 + \gamma t) > c = d/a$, $0 \leqq t < t_1$.

When the water supply to the turbines is reduced to a comparatively small value so that $d_1 \gg d_2$, the initial conditions are $u(0) = d_1/a$, $\dot{u}(0) = 0$, and we take $c = d_2/a$. Thus,

$$A_0 = \frac{(d_1 - d_2)/a}{\{1 - [4\kappa(d_1 - d_2)/3\pi\omega a]^2\}^{\frac{1}{2}}}, \tag{26}$$

and

$$u = \frac{A_0}{1 + \gamma t} \cos \left(\omega t - \tan^{-1} \frac{\gamma}{\omega}\right) + \frac{d_2}{a}, \tag{27}$$

provided that $A_0/(1 + \gamma t) > d_2/a$, and $0 \leqq t < t_1 = (3\pi/4\kappa)(a/d_2 - 1/A_0)$. When $t > t_1$, the solution has the form at (14), *viz.*,

$$u = A_1 e^{-\kappa c t} \cos (\omega t - \phi_1) + c, \tag{28}$$

where A_1 and ϕ_1 may be determined from the values of u and $\dot{u}$ at $t = t_1$. Then by (25)

$$u(t_1) = \frac{A_0}{1 + \gamma t_1} \cos (\omega t_1 - \phi_0) + c, \tag{29}$$

and

$$\dot{u}(t_1) = -\frac{A_0}{1 + \gamma t_1} \left[\frac{\gamma \cos (\omega t_1 - \phi_0)}{1 + \gamma t_1} + \omega \sin (\omega t_1 - \phi_0)\right]. \tag{30}$$

From (28), (29), and (30), we find that

$$A_1 = \frac{A_0 e^{\kappa c t_1}}{1 + \gamma t_1} \frac{\cos (\omega t_1 - \phi_0)}{\cos (\omega t_1 - \phi_1)}, \tag{31}$$

and

$$\phi_1 = \omega t_1 + \tan^{-1} \left[\frac{\kappa c}{\omega} - \frac{\gamma}{\omega(1 + \gamma t_1)} - \tan (\omega t_1 - \phi_0)\right]. \tag{32}$$

$|A| > c = 0$, $t \geqq 0$. Suppose that when the velocity is d_1/a, the water supply to the turbines is stopped completely. The initial conditions are now $u(0) = d_1/a$, $\dot{u}(0) = 0$, with $c = 0$, and the required results may be obtained from (26), (27) by putting $d_2 = 0$. Thus

$$A_0 = \frac{d_1}{a} \bigg/ \left[1 - \left(\frac{4\kappa d_1}{3\pi\omega a}\right)^2\right]^{\frac{1}{2}}, \tag{33}$$

and the velocity of the water in the tunnel is

$$u = \frac{A_0}{1 + \gamma t} \cos \left(\omega t - \tan^{-1} \frac{\gamma}{\omega}\right). \tag{34}$$

Comparison of (34) with (19) shows that the decay of the oscillation is more

rapid in the latter case, owing to the respective damping factors being $1/(1 + \gamma t)$ and $e^{-\kappa d_2 t/a}$. A flow of water entails $d_2 > 0$, so that in (2) the respective damping coefficients for $d_2 > 0$ and $d_2 = 0$ are $2\kappa|x + c|$ and $2\kappa|x|$, the former being greater than the latter.

6. The equation $\ddot{y} + 2\kappa y\dot{y} + ay = 0$ $(a,\kappa > 0)$. This is (1) with $d = 0$ and y written for $|y|$, so that the damping coefficient changes sign with y. If the equation is that for a dynamical or other physical system, when $y > 0$ some of the energy in the system is dissipated, whereas when $y < 0$ energy is supplied to the system from an external source. In the case of an electrical circuit, there would be a negative-resistance effect as with an electron-tube oscillator.

Taking $v = dy/dt$, $v' = dv/dy$, the above equation may be written

$$v\,dv + 2\kappa y(b + v)\,dy = 0 \tag{35}$$

with $b = a/2\kappa$. Integrating (35) gives

$$v - b \ln (b + v) + \kappa y^2 = A, \text{ a const.,} \tag{36}$$

provided $v > -b$. When $v < -b$, the appropriate result is

$$v - b \ln [-(b + v)] + \kappa y^2 = B, \text{ a const.} \tag{37}$$

If the initial conditions for (36) are $y = y_0 > 0$, $v = v_0 = 0$, we obtain

$$v - b \ln \frac{b + v}{b} + \kappa(y^2 - y_0^2) = 0, \tag{38}$$

and

$$y = \pm\left[\frac{1}{\kappa}\left(-v + b \ln \frac{b + v}{b}\right) + y_0^2\right]^{\frac{1}{2}}; \tag{39}$$

so the (v,y)-curves are symmetrical about the v-axis, when $v > -b$. Also the points $y = \pm y_0$, $v = 0$ are extrema on the y-axis, while on the v-axis, where $y = 0$, there are two values, one positive and the other negative $(> -b)$, which satisfy

$$v - b \ln \frac{b + v}{b} - \kappa y_0^2 = 0. \tag{40}$$

These values, of which the positive one is the greater numerically, are extrema on the v-axis.

From (35)

$$v' = -y(a + 2\kappa v)/v, \tag{41}$$

and the origin $y = v = 0$ is either a center or a spiral point. Equation (41) shows that the (v,y)-curves intersect the axes orthogonally. If for any v the sign of y is changed, so also is that of v'. Also, the slopes of the (v,y)-curves are negative in the first and third quadrants, positive in the second and fourth quadrants, while by (39) the curves are symmetrical about the v-axis. From these deductions together with the existence of extrema on the two axes, we

conclude that the origin is a center, which entails stable periodic motion, provided $v > -b$.

Formula (38) may be written

$$v - b \ln\left(1 + \frac{v}{b}\right) + \kappa y^2 = \kappa y_0^2. \tag{42}$$

If $|v/b| < 1$ is small enough, (42) takes the approximate form

$$\frac{v^2}{2b} + \kappa y^2 = \kappa y_0^2, \qquad \text{or} \qquad \frac{y^2}{y_0^2} + \frac{v^2}{a y_0^2} = 1, \tag{43}$$

which is the equation for an ellipse with the origin as center and semiaxes $y_0, a^{\frac{1}{2}} y_0$.

The (v,y)-curves for (37) entail v and $v_0 < -b$. Thus with initial conditions $y = y_0$, $v = v_0 < -b$, we have

$$v = v_0 - b \ln \frac{b + v_0}{b + v} + \kappa(y_0^2 - y^2), \tag{44}$$

so that

$$y = \pm \left[\frac{1}{\kappa}\left((v_0 - v) - b \ln \frac{b + v_0}{b + v}\right) + y_0^2\right]^{\frac{1}{2}}, \tag{45}$$

which indicates that when $v < -b$, the (v,y)-curves are symmetrical about the v-axis. Also, it follows from (44) that, with $v_0 < -b$, $v \to -\infty$ as $y \to -\infty$, whatever the value of y_0. Hence, under the foregoing initial conditions, the motion of a physical system which obeys the differential equation

$$\ddot{y} + 2\kappa y \dot{y} + a y = 0 \qquad (a, \kappa > 0)$$

is unstable, and $y \to -\infty$ as $t \to +\infty$.

The forms of the (v,y)-curves corresponding to the stable $(v > -b)$ and unstable $(v < -b)$ cases are illustrated in Fig. 5.

7. Physical considerations. In the schematic diagram, Fig. 6, s is a uniform coil spring of negligible mass, whose force-displacement relationship is $f = ay$, $a > 0$. m is a unit mass constrained by frictionless guides to move along the axis of the spring. The damping device is such that when m is moving to the right[3] the damping coefficient is $2\kappa y > 0$, while for motion to the left it is $2\kappa y < 0$. Let m be displaced to the right a distance y_0, thereby extending s by this amount. On m being released, its velocity will increase until the central position $y = 0$ is reached. During the corresponding time interval τ_0, there will be energy loss $2\kappa \int_0^{\tau_0} y v^2\, dt$, while, during the succeeding interval τ_0, this energy will have been restored; so m reaches the extreme position $y = -y_0$. On the journey from $y = -y_0$ to 0, $2\kappa y < 0$, so that energy will be supplied to the system over a time interval τ_1, say. Consequently the velocity of m in the central position will exceed that when it has moved there from $y = y_0$. The kinetic energy of m is now greater than the potential

[3] Of the central position.

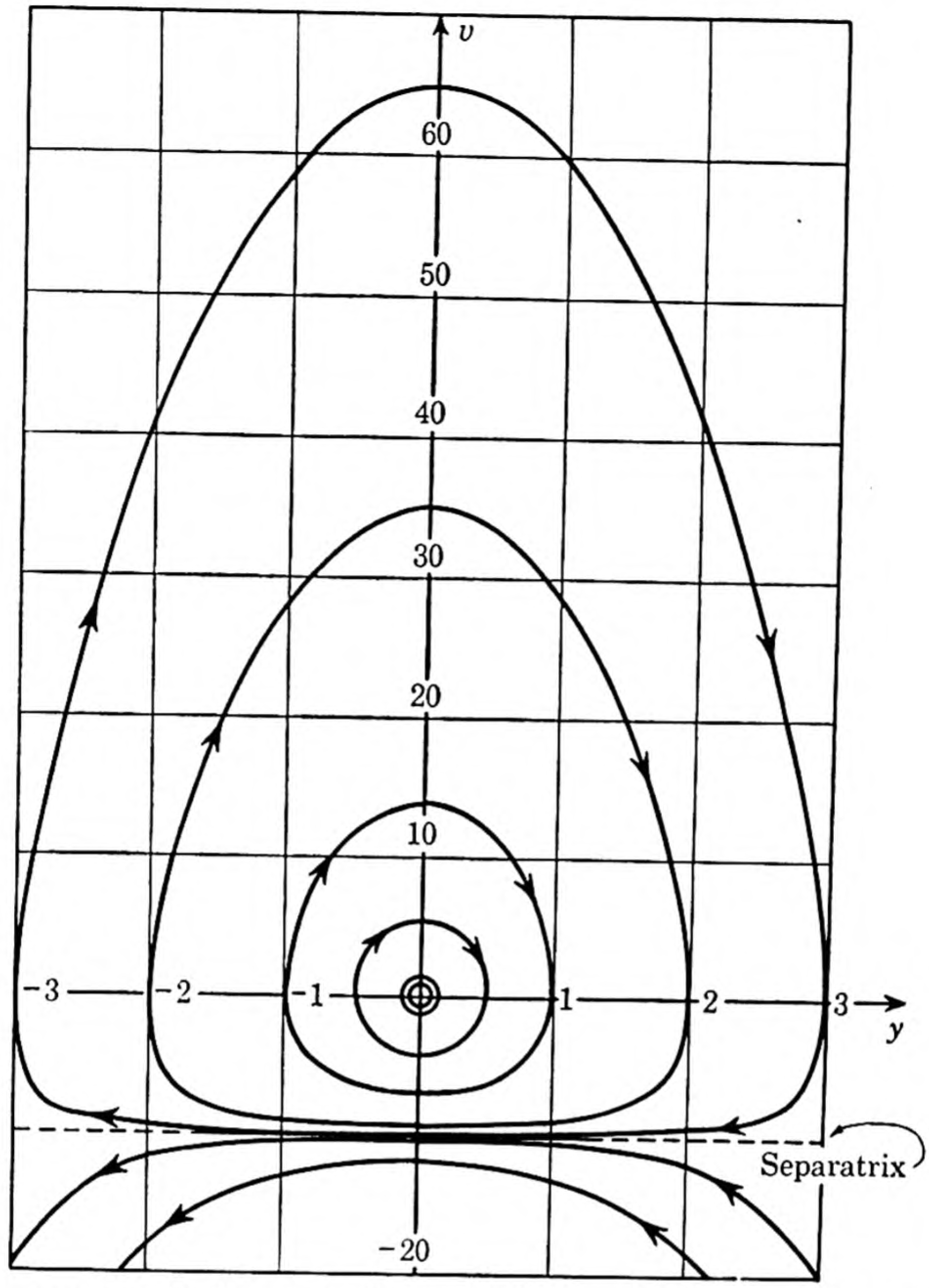


FIG. 5. Integral curves $(y,\dot{y})$ for $\ddot{y} + 10y\dot{y} + 100y = 0$, showing stable (above the separatrix) and unstable regions (below the separatrix). The origin is a center. $-a/2\kappa = -10$.

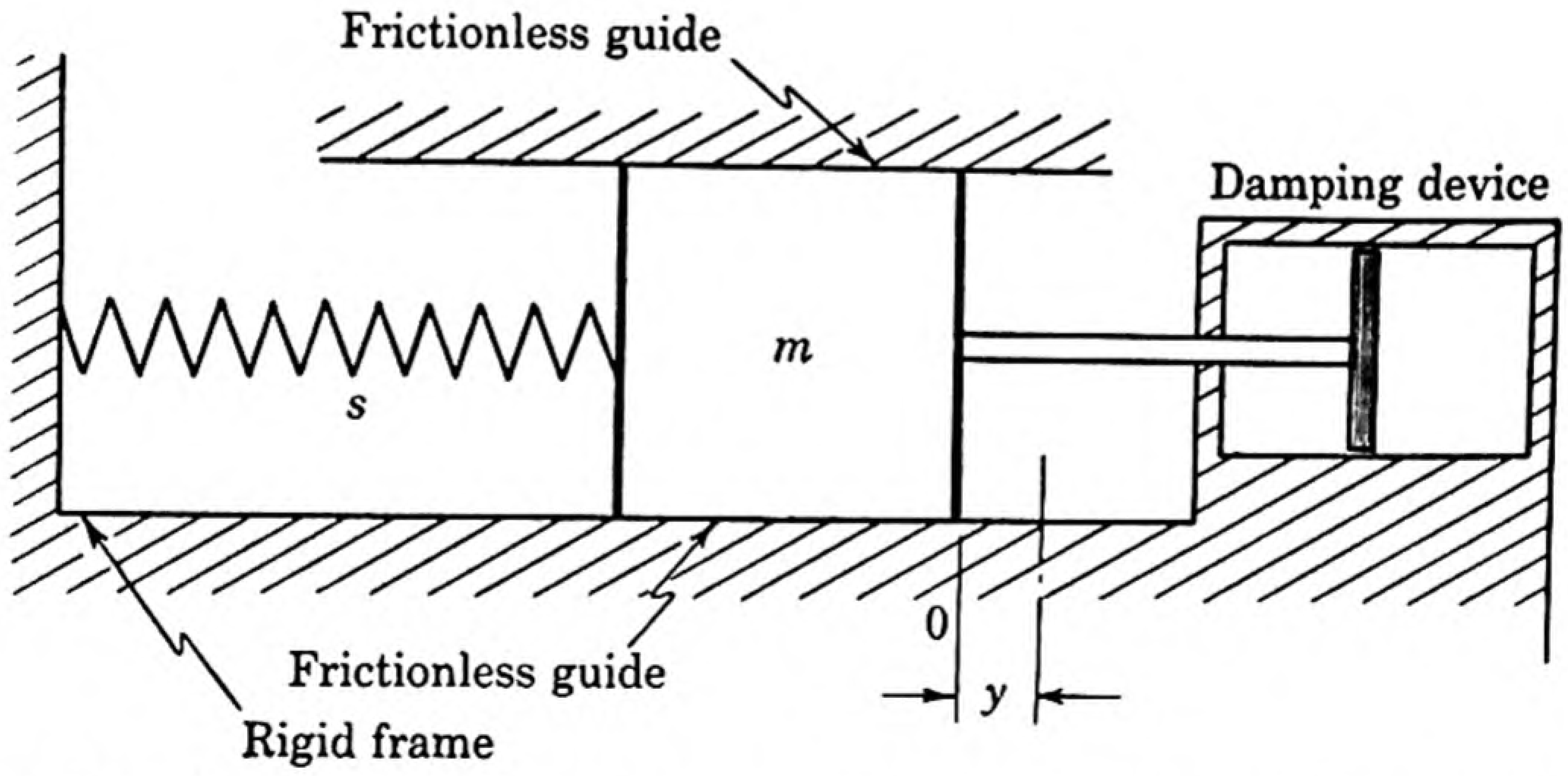


FIG. 6. Schematic diagram for mass-spring system having positive-negative damping device.

energy at either $y = \pm y_0$ by an amount $2\kappa \int_0^{\tau_1} yv^2\, dt$. During the succeeding interval τ_1, the damping device causes dissipation, and the whole of the energy gained in the previous τ_1 interval has been lost when m reaches $y = y_0$.

If y_0 is large enough, the (v,y)-curve in the third and fourth quadrants will be sensibly rectangular, the velocity over most of the path being approximately $v = -a/2\kappa$. Thus $2\tau_0 \simeq 4\kappa y_0/a$, which is the approximate time of travel from y_0 to $-y_0$, so $2\tau_0 \to \infty$ as $y_0 \to \infty$.

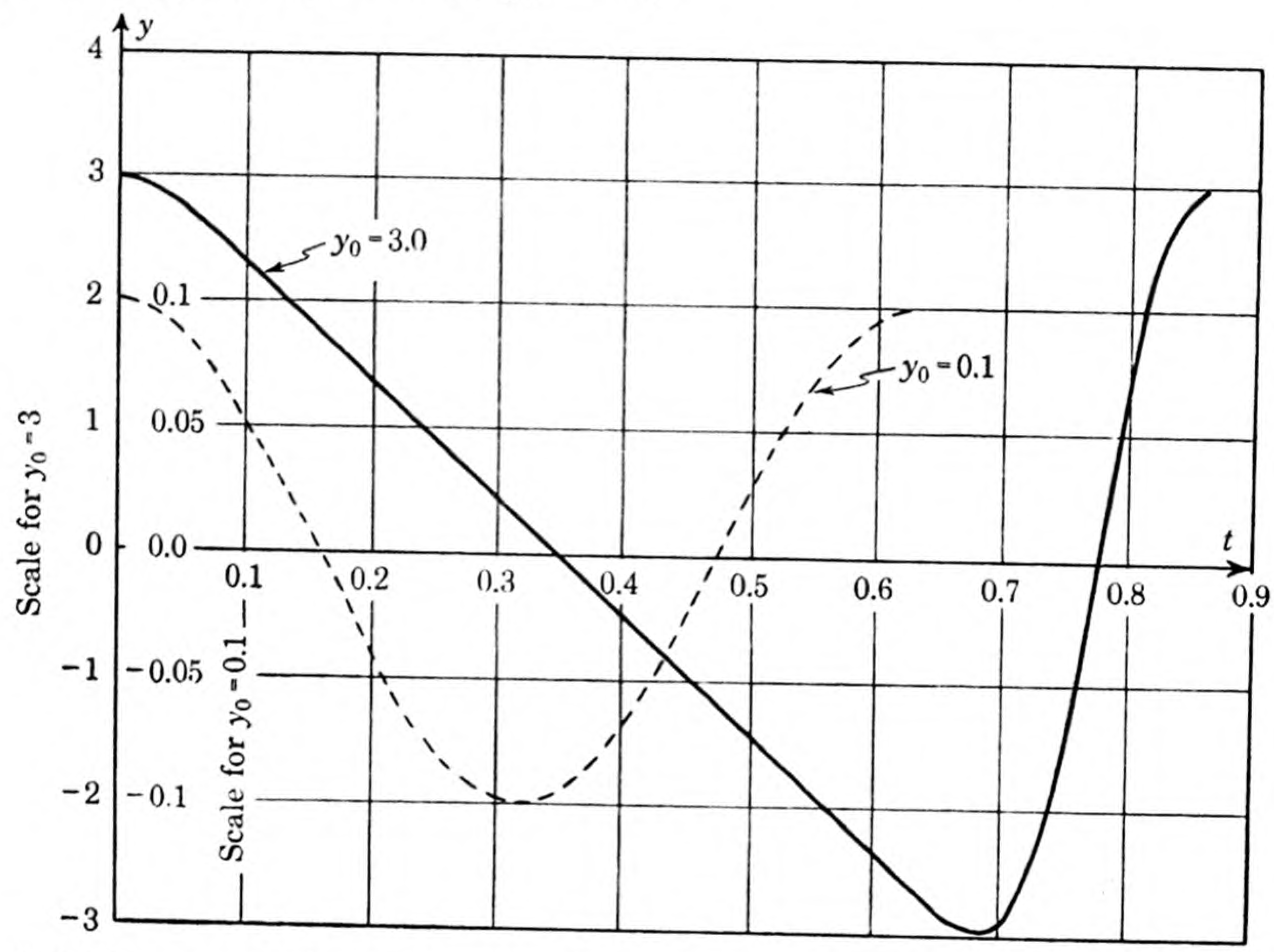


FIG. 7. Solution curves (y,t), for $\ddot{y} + 10y\dot{y} + 100y = 0$. Solid curve is for initial conditions $y(0) = y_0 = 3$, $\dot{y}(0) = 0$; for broken curve $y(0) = y_0 = 0.1$, $\dot{y}(0) = 0$.

In the first and second quadrants, so long as v is not too small, the logarithmic term may be neglected in (38), and we get

$$v \simeq \kappa(y_0^2 - y^2) = \kappa y_0^2\left[1 - \left(\frac{y}{y_0}\right)^2\right]. \tag{46}$$

The time interval corresponding to the movement of m over any range $(-y_1,y_1)$, where $|y_1| < |y_0|$, for which (46) is valid, is given by $t = 2\int_0^{y_1} dy/v$, and t decreases with increase in y_0. The time intervals for the ranges $(-y_0,-y_1)$ and (y_1,y_0) are correspondingly small. Hence the periodic time $2(\tau_0 + \tau_1)$ increases indefinitely with increase in y_0. For large finite values of the latter, the (y,t)-curve has the form of a saw-tooth wave with rounded corners, as exemplified in Fig. 7.

To explain the (v,y)-curves below the line $v = -a/2\kappa$, consider the initial conditions to be such that when $y = y_0$, m has a velocity $v_0 < -a/2\kappa$. In the range $y < 0$, the control ay is inadequate to prevent v increasing (negatively)

under the influence of negative damping, so that both y and $v \to -\infty$ as $t \to +\infty$.

8. Approximate solution of $\ddot{y} + 2\kappa y\dot{y} + ay = 0$, **$y$ small,** $v > -a/2\kappa$. If κ is relatively small, the method of perturbation may be used, since the solution is periodic. Thus we assume that

$$y = y_0 + \kappa y_1 + \kappa^2 y_2 + \cdots, \tag{47}$$

and

$$a = \omega_0^2 + \kappa\omega_1^2 + \kappa^2\omega_2^2 + \cdots, \tag{48}$$

where $y_0, y_1, y_2, \cdots$ are twice differentiable functions to be determined, while $\omega_0^2, \omega_1^2, \cdots$ are evaluable constants, in virtue of the periodic nature of the solution. Proceeding in known manner and writing $\psi_0 = \omega_0 t$, to the second order in κ, we find that, with initial conditions $y(0) = y_0$, $\dot{y}(0) = 0$,

$$y = y_0\left(1 - \frac{23\kappa^2 y_0^2}{72\omega_0^2}\right)\cos\psi_0 + \frac{\kappa y_0^2}{3\omega_0}(2\sin\psi_0 - \sin 2\psi_0) + \frac{\kappa^2 y_0^3}{72\omega_0^2}(32\cos 2\psi_0 - 9\cos 3\psi_0), \tag{49}$$

where $\omega_0^2 = a - (\kappa^2 y_0^2/3)$, and we assume that $\kappa^2 y_0^2/3 \ll a$.

For initial conditions $y(0) = 0$, $\dot{y}(0) = \omega_0 y_0$, to order 2 in κ the solution is

$$y = y_0\left(1 - \frac{2\kappa y_0}{3\omega_0} + \frac{37\kappa^2 y_0^2}{72\omega_0^2}\right)\sin\psi_0 + \frac{\kappa y_0^2}{3\omega_0}\left(1 - \frac{4\kappa y_0}{3\omega_0}\right)\sin 2\psi_0 + \frac{\kappa^2 y_0^3}{8\omega_0^2}\sin 3\psi_0, \tag{50}$$

with ω_0^2 as above.

In these instances the (v,y)-curves would be approximately elliptical in shape with the origin as center, particularly so if $a/2\kappa$ were large. The periodic time of the motion is

$$\tau = \frac{2\pi}{\omega_0} \simeq \frac{2\pi(1 + \kappa^2 y_0^2/6a)}{a^{\frac{1}{2}}}, \tag{51}$$

so τ increases with increase in y_0. This feature persists however large y_0 may be, as was demonstrated in Sec. 7.

9. Consideration of $\ddot{y} + 2\kappa y\dot{y} + ay = d > 0$. Writing $y = x + c$, we obtain

$$\ddot{x} + 2\kappa(x + c)\dot{x} + ax = 0, \tag{52}$$

provided $c = d/a$. The paragraph below (6) shows that if $\kappa^2 c^2 < a$, the point $y = c$, $v = 0$ is a stable spiral point, while if $\kappa^2 c^2 \geqq a$, it is a stable node which the (v,y)-curve enters as $t \to +\infty$. Thus the respective motions corresponding to these conditions are a damped oscillation and aperiodicity (absence of oscillation). When $\kappa^2 c^2 \ll a$—the values of $y(0)$, $\dot{y}(0)$, and κ not being too

large—in the first approximation the solution is given by (14), so that with y for u we have

$$y = A_0 e^{-\kappa c t} \cos (\omega t - \phi_0) + \frac{d}{a}. \tag{53}$$

If the initial conditions are $y = y_0$, $v = 0$, then

$$y = \left(y_0 - \frac{d}{a}\right)\left[1 + \left(\frac{\kappa d}{a\omega}\right)^2\right]^{\frac{1}{2}} e^{-\kappa c t} \cos (\omega t - \phi_0) + \frac{d}{a}, \tag{54}$$

with $\phi_0 = \tan^{-1} (\kappa d/\omega a)$.

BIBLIOGRAPHY

1. H. Poincaré, J. Math. Pures Appl. vol. 8 (1882) p. 251.
———, *Les méthodes nouvelles de la mécanique céleste*, vol. 1, Paris, 1892.
2. N. Kryloff and N. Bogoliuboff, *Introduction to non-linear mechanics*, translation from the Russian by S. Lefschetz, Princeton, 1943.
3. A. M. Binnie, Proc. Camb. Phil. Soc. vol. 42 (1946), p. 156.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

ACOUSTIC RADIATION PRESSURE ON A CIRCULAR DISK[1]

BY

HAROLD LEVINE

1. Introduction. The radiation pressure experienced by material objects in a wave field represents an important acoustical effect. It is therefore of interest to have the related theory available for a variety of experimental conditions.

A calculation of radiation pressure on a plane circular disk of rigid material by King [1] furnishes the inspiration for this paper. The disk is exposed to small-amplitude plane progressive or stationary harmonic waves with equiphase surfaces parallel to its plane and performs oscillations along the normal direction. In the analysis of King, both inertial and diffraction effects are taken into account approximately, the latter only for wavelengths large compared with the disk radius. We present here another method of calculating the radiation pressure which is capable of high accuracy for all wavelengths.

King's procedure employs integral representations of the wave function and leads to a pair of dual integral equations, one in the domain of the disk and the other in the remainder of its plane. An approximate solution of these equations is effected by appealing to the known solution of related frequency-independent or static integral equations. This process furnishes the first few terms in the low-frequency expansion of the solution, with succeeding terms obtainable after tedious computations. Estimates of the radiation pressure are thereby restricted to low frequencies or long wavelengths.

The scheme proposed here involves a single integral equation which determines the difference in velocity potential on opposite disk faces. From the integral equation and its connection with the normal radiation pressure, we construct an expression which is stationary for variations about the correct solution and has a stationary value equal to that of the radiation pressure. The stationary form has the practical merit of subordinating difficulties associated with the construction of an exact or generally accurate solution to the integral equation. We shall illustrate this approach by extending the results of King to include the high-frequency limit.

2. Formulation. At equilibrium we locate the disk in the plane $z = 0$. When the incident radiation field has axial symmetry with respect to z, a circular disk will carry out oscillations perpendicular to its plane, with no tendency for angular deflections. If $\psi(\mathbf{r})$ denote the velocity potential of a field possessing harmonic time dependence $\exp(-i\omega t)$, the conditions of the problem are met by constructing a solution of the wave equation

[1] This work was performed on a National Bureau of Standards contract with the University of California, Los Angeles, and was sponsored (in part) by the Office of Naval Research, U.S. Navy.

$$(1) \qquad (\nabla^2 + k^2)\psi(\mathbf{r}) = 0, \qquad k = \frac{\omega}{c}$$

(c denotes the sound velocity) which incorporates the prescribed form of incident radiation and has normal derivative

$$(2) \qquad -\frac{\partial}{\partial z}\psi(\mathbf{r}) = \dot{\zeta}$$

on the disk, $\dot{\zeta}$ being the disk velocity. The latter quantity is taken from the dynamical equation for the disk, whose motion is ascribed to the first-order pressure difference on opposite faces. Employing the scalar Green's function

$$(3) \qquad G(\mathbf{r},\mathbf{r}') = \frac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|} = G(\mathbf{r}',\mathbf{r}),$$

which satisfies the inhomogeneous differential equation

$$(4) \qquad (\nabla^2 + k^2)G(\mathbf{r},\mathbf{r}') = -\delta(\mathbf{r}-\mathbf{r}'),$$

a representation of the velocity potential deduced from Green's integral theorem is

$$(5) \qquad \psi(\mathbf{r}) = \psi^{inc}(\mathbf{r}) + \int [G(\mathbf{r}',\mathbf{r})\,\partial_{n'}\psi(\mathbf{r}') - \psi(\mathbf{r}')\,\partial_{n'}G(\mathbf{r}',\mathbf{r})]\,ds',$$

where the integral extends over both disk faces. Since the derivatives in (5) refer to the inward normal and the disk moves rigidly, $\partial_{n'}\psi$ has equal and opposite values for any pair of points on the two faces. Then the first term of the integral vanishes, giving

$$(6) \qquad \psi(\mathbf{r}) = \psi^{inc}(\mathbf{r}) - \int_A \Psi(\mathbf{r}')\,\partial_{z'}G(x',y',0,\mathbf{r})\,ds',$$

where

$$(7) \qquad \Psi(\mathbf{r}) = \psi(x,y,z\to 0-) - \psi(x,y,z\to 0+) \qquad (x,y \text{ in } A)$$

and A denotes the disk area.

With the help of the Green's function Fourier-integral representation

$$(8) \quad G(\mathbf{r},\mathbf{r}') = \frac{i}{8\pi^2} \int_{-\infty}^{\infty} \frac{\exp\{i[k_x(x-x') + k_y(y-y') + \sqrt{k^2-k_x^2-k_y^2}\,|z-z'|]\}}{\sqrt{k^2-k_x^2-k_y^2}}\,dk_x\,dk_y$$

and the Dirac delta-function relations

$$(9) \qquad \begin{aligned} \delta(x-x')\,\delta(y-y') &= (2\pi)^{-2}\int_{-\infty}^{\infty} \exp\{i[k_x(x-x') + k_y(y-y')]\}\,dk_x\,dk_y \\ \int_A f(x',y')\,\delta(x-x')\,\delta(y-y')\,dx'\,dy' &= f(x,y), \qquad (x,y \text{ in } A), \end{aligned}$$

it follows from (6) that

$$(10) \qquad \begin{aligned} \psi(x,y,z\to 0-) &= \psi^{inc}(x,y,z\to 0-) + \tfrac{1}{2}\Psi(\mathbf{r}), \\ \psi(x,y,z\to 0+) &= \psi^{inc}(x,y,z\to 0+) - \tfrac{1}{2}\Psi(\mathbf{r}). \end{aligned}$$

On subtracting the latter equations and noting that

$$\psi^{inc}(x,y,z \to 0-) = \psi^{inc}(x,y,z \to 0+) = \psi^{inc}(x,y,0),$$

we obtain the identity

$$\Psi(\mathbf{r}) = \Psi(\mathbf{r}),$$

whereas addition yields the new result

(11) $$\psi(x,y,z \to 0-) + \psi(x,y,z \to 0+) = 2\psi^{inc}(x,y,0).$$

For the progressive incident wave

(12) $$\psi^{inc}(\mathbf{r}) = e^{ikz},$$

we satisfy the boundary condition (2) if

(13) $$-\dot{\zeta} = ik - \int_A \Psi(\mathbf{r}')K(\mathbf{r}',\mathbf{r})\, ds' \qquad (\mathbf{r} \text{ on } A)$$

and

(14) $$K(\mathbf{r},\mathbf{r}') = \partial_z\partial_{z'} G(x,y,0,x',y',0) = K(\mathbf{r}',\mathbf{r}).$$

To write a dynamical equation for the disk requires information about the pressure in the medium. It is well known that with inclusion of quadratic or second-order terms in the velocity potential the deviation of pressure p from its equilibrium value p_0 takes the form [1]

(15) $$\Delta p = p - p_0 = \rho_0\dot{\psi} + \frac{\rho_0}{2c^2}\dot{\psi}^2 - \frac{\rho_0}{2}(\nabla\psi)^2$$

where ρ_0 is the equilibrium density of the medium. Using only the linear term in (15), the equation of motion for a disk of mass m_0 admits the first integral

(16) $$m_0\dot{\zeta} = \rho_0 \int_A \Psi(\mathbf{r})\, ds.$$

Thus, by eliminating $\dot{\zeta}$ between (13) and (16), we arrive at an integral equation for the function $\Psi(\mathbf{r})$,

(17) $$\int_A \Psi(\mathbf{r}')K(\mathbf{r}',\mathbf{r})\, ds' - \frac{\rho_0}{m_0}\int_A \Psi(\mathbf{r})\, ds = ik, \qquad (\mathbf{r} \text{ on } A),$$

which is specialized to the rigidly fixed disk if $m_0 \to \infty$.

We next inquire after the time-average radiation pressure on the disk in the normal or z direction. The linear term in (15) evidently does not contribute to the result, and likewise for the final term, since

$$\int_A [(\nabla\psi(x,y,z \to 0-))^2 - (\nabla\psi(x,y,z \to 0+))^2]\, ds$$
$$= 2\int_A [\partial_x\Psi(\mathbf{r})\,\partial_x\psi^{inc}(\mathbf{r}) + \partial_y\Psi(\mathbf{r})\,\partial_y\psi^{inc}(\mathbf{r})]\, ds = 0,$$

in consequence of (2), (7), (11), and (12). Finally, the normal pressure becomes

$$P = \frac{1}{2}\frac{\rho_0}{c^2}\int_A [(\dot\psi(x,y,z \to 0-))^2 - (\dot\psi(x,y,z \to 0+))^2]\, ds \tag{18}$$
$$= \frac{\rho_0}{c^2}\dot\psi^{inc}\int_A \dot\Psi(\mathbf{r})\, ds,$$

or, on taking the time average,

$$\bar P = \frac{\rho_0 k^2}{2}\,\mathrm{Re}\int_A \Psi(\mathbf{r})\, ds. \tag{19}$$

A different characterization of the time-average radiation pressure is obtained by employing the integral equation (17) for $\Psi(\mathbf{r})$. On multiplication of the latter equation by $\Psi(\mathbf{r})$ and subsequent integration over A, there results

$$\int_A \Psi(\mathbf{r})K(\mathbf{r},\mathbf{r}')\Psi(\mathbf{r}')\, ds\, ds' - \frac{\rho_0}{m_0}\left(\int_A \Psi(\mathbf{r})\, ds\right)^2 = ik\int_A \Psi(\mathbf{r})\, ds, \tag{20}$$

whence

$$\int_A \Psi(\mathbf{r})K(\mathbf{r},\mathbf{r}')\Psi(\mathbf{r}')\, ds\, ds' - \frac{\rho_0}{m_0}\left(\int_A \Psi(\mathbf{r})\, ds\right)^2 - 2ik\int_A \Psi(\mathbf{r})\, ds \tag{21}$$
$$= -ik\int_A \Psi(\mathbf{r})\, ds.$$

One easily verifies that the left-hand member of (21) is stationary for variations of $\Psi(\mathbf{r})$ about the solution to the integral equation (17), with the right-hand member as its stationary value. Moreover, by utilizing the relation (19), we deduce that the time-average radiation pressure is the stationary value of the functional

$$\bar P(= \bar P(\Psi)) = -\frac{1}{2}\rho_0 k\,\mathrm{Im}\left[\int_A \Psi(\mathbf{r})K(\mathbf{r},\mathbf{r}')\Psi(\mathbf{r}')\, ds\, ds' - \frac{\rho_0}{m_0}\left(\int_A \Psi(\mathbf{r})\, ds\right)^2 - 2ik\int_A \Psi(\mathbf{r})\, ds\right]; \tag{22}$$

this feature imparts practical significance to (22) as a means of approximating the radiation pressure. In the latter connection, it is useful to have another form of (22),

$$\bar P = -\frac{1}{2}\rho_0 k^3\,\mathrm{Im}\frac{\left(\int_A \Psi(\mathbf{r})\, ds\right)^2}{\int_A \Psi(\mathbf{r})K(\mathbf{r},\mathbf{r}')\Psi(\mathbf{r}')\, ds\, ds' - \frac{\rho_0}{m_0}\left(\int_A \Psi(\mathbf{r})\, ds\right)^2}, \tag{23}$$

which is independent of the scale for $\Psi(\mathbf{r})$.

Before applying (23) to the specific case of a circular disk, we may remark on some other general features. Thus, when the disk has infinite inertia ($m_0 = \infty$), the radiation pressure and plane-wave scattering cross section at normal incidence, σ_{sc}, of the disk are simply related by

(24) $$\bar{P} = \tfrac{1}{2}\rho_0 k^2 \sigma_{sc}, \qquad (m_0 = \infty),$$

or

(25) $$\frac{\bar{P}}{\bar{W}^{inc}} = \sigma_{sc},$$

where

(26) $$\bar{W}^{inc} = \tfrac{1}{2}\rho_0 k^2$$

is the time-average energy density for the incident wave (12). The relation (24) is substantiated on exhibiting a stationary scale independent form of the cross section in terms of $\Psi(\mathbf{r})$, for details of which we refer to another paper [2].

In the limit of extremely short wavelengths, or very high frequencies, the normal scattering cross section for a rigid disk approaches a value of twice its area A, and consequently

(27) $$\bar{P} \simeq 2\bar{W}^{inc}A, \qquad (k \to \infty).$$

This result can be understood in terms of the momentum transfer which accompanies reflection of the incident plane wave at the disk, with the factor 2 specific for unit reflection coefficient at a rigid surface. Even if the disk has a finite mass m_0, the asymptotic relation (27) holds, for the first term in the denominator of (23) then dominates the second.

At low frequencies, examination of the integral equation (17) shows that $\Psi(\mathbf{r})$ is wholly imaginary and (except for a scale factor) independent of frequency. With this knowledge, and use of the expansion

(28) $$\begin{aligned} K(\mathbf{r},\mathbf{r}') &= K_s(\mathbf{r},\mathbf{r}') + \frac{ik^3}{12\pi} + \cdots, \qquad (k \to 0), \\ K_s(\mathbf{r},\mathbf{r}') &= \partial_z \partial_{z'} \left.\frac{1}{|\mathbf{r} - \mathbf{r}'|}\right|_{z=z'=0}, \end{aligned}$$

it follows by (23) that

(29) $$\bar{P} = \frac{1}{2}\rho_0 k^2 \sigma_{sc} \frac{1}{[1 + (m_1/m_0)]^2}, \qquad (k \to 0),$$

where

(30) $$m_1 = \rho_0 \frac{\left(\int_A \Psi(\mathbf{r})\, ds\right)^2}{\int_A \Psi(\mathbf{r}) K_s(\mathbf{r},\mathbf{r}')\Psi(\mathbf{r}')\, ds\, ds'},$$

and

(31) $$\sigma_{sc} = \frac{k^4}{12\pi}\frac{m_1^2}{\rho_0^2}, \qquad (k \to 0).$$

The inertial effect is thus isolated in a single factor and characterized by the quantity m_1, the so-called hydrodynamic mass of the disk.

A final remark in connection with the actual use of (23) is appropriate; as $K(\mathbf{r},\mathbf{r}')$ is a highly singular function, nonvanishing estimates of the radia-

tion pressure are obtained only for the restricted class of trial functions $\Psi(\mathbf{r})$ that vanish on the rim of the disk.

3. Application. In regard to a circular disk with radius R, the function $\Psi(r)$ may be adequately described by linear combinations of the primitive form $[1 - (r^2/R^2)]^{n+\frac{1}{2}}$, $n = 0, 1, \cdots$, where r denotes the radial coordinate referred to the disk center. Thus, as a first approximation, we take

$$\Psi_1(r) = \left(1 - \frac{r^2}{R^2}\right)^{\frac{1}{2}}, \tag{32}$$

in (23), and avail ourselves of the results in [2] to deduce that

$$\bar{P}_1 = \frac{4}{9}\rho_0(kR)^3 \frac{I_{11}(kR)}{[R_{11}(kR) + \frac{8}{9}(\rho_0 R^3/m_0)]^2 + [I_{11}(kR)]^2}, \tag{33}$$

with

$$I_{11}(\alpha) = \frac{\alpha}{2\pi} - \frac{1}{4\pi\alpha} + \frac{1}{8\alpha^2} S_0(2\alpha) + \frac{1}{4\alpha} S_1(2\alpha) - \frac{1}{4\alpha}\left(1 + \frac{1}{4\alpha^2}\right)\int_0^{2\alpha} S_0(t)\,dt,$$

and

$$R_{11}(\alpha) = \frac{1}{4\alpha}\left(1 + \frac{1}{4\alpha^2}\right)\int_0^{2\alpha} J_0(t)\,dt - \frac{1}{8\alpha^2} J_0(2\alpha) - \frac{1}{4\alpha} J_1(2\alpha).$$

Here J_0, J_1 and S_0, S_1 denote Bessel and Struve functions of order zero and 1, respectively. Using well-known properties of these functions, it follows that, for small values of kR,

$$I_{11}(kR) = \frac{2}{27\pi}(kR)^3, \qquad R_{11}(kR) = \frac{1}{3},$$

whence

$$\bar{P}_1 = \frac{8}{27\pi}\rho_0(kR)^6 \frac{1}{[1 + \frac{8}{3}\rho_0(R^3/m_0)]^2}, \qquad (kR \ll 1). \tag{34}$$

A comparison of (29) and (34) establishes

$$m_1 = \tfrac{8}{3}\rho_0 R^3, \tag{35}$$

as the hydrodynamic mass for the circular disk with radius R.

The result (34) also emerges from King's calculations, which, however, disagree slightly in the succeeding frequency correction. Thus, with two terms for the estimate of $R_{11}(kR)$, $kR \ll 1$, equation (23) yields

$$\bar{P}_1 = \frac{8}{27\pi}\rho_0(kR)^6 \frac{1 + \frac{8}{25}(kR)^2}{[1 + \frac{8}{3}\rho_0(R^3/m_0)(1 + \frac{1}{5}(kR)^2)]^2}, \tag{36}$$

while King gives the (incorrect) numerical coefficient $\frac{1}{4}$ instead of $\frac{8}{25}$.

At very high frequencies, where King's procedure becomes impractical, the asymptotic behavior of $\bar{P}_1$,

$$\bar{P}_1 \simeq \tfrac{8}{9}\pi\rho_0(kR)^2, \qquad (kR \gg 1), \tag{37}$$

is easily derived, since then

$$I_{11}(kR) \simeq \frac{kR}{2\pi}, \qquad R_{11}(kR) \simeq 0.$$

To ascertain the approximate nature of (37), we rewrite it in the form

$$\bar{P}_1 \simeq \tfrac{1}{2}\rho_0 k^2(\tfrac{16}{9}\pi R^2), \tag{38}$$

which permits ready comparison with (26), (27). This shows that (38) deviates from the correct result by a numerical factor in the scattering cross section, having $\frac{16}{9}$ in place of 2. For intermediate values of kR, numerical values of $\bar{P}_1$ are obtained from (23) without difficulty.

An improvement in these results is achieved by use of the trial function

$$\Psi_2(r) = \left(1 - \frac{r^2}{R^2}\right)^{\frac{1}{2}} + c\left(1 - \frac{r^2}{R^2}\right)^{\frac{3}{2}}, \tag{39}$$

where c is a variational parameter, that can be specified by the stationary condition

$$\frac{\partial \bar{P}_2}{\partial c} = 0.$$

In this way we derive an estimate $\bar{P}_2$ of the radiation pressure which agrees precisely at low frequencies with $\bar{P}_1$ and is generally more accurate at higher frequencies, attaining asymptotically the form

$$\bar{P}_2 \simeq \tfrac{1}{2}\rho_0 k^2(\tfrac{48}{25}\pi R^2) = \tfrac{24}{25}\pi\rho_0(kR)^2, \qquad (kR \gg 1). \tag{40}$$

The case of excitation by stationary plane waves is treated analogously.

Bibliography

1. L. V. King, Proc. Roy. Soc. London Ser. A vol. 153 (1935) p. 1.
2. H. Levine and J. Schwinger, Phys. Rev. vol. 74 (1948) p. 958.

University of California,
Los Angeles, Calif.
(on Leave from Harvard University, Cambridge, Mass.)

INFINITE MATRICES ASSOCIATED WITH A DIFFRACTION PROBLEM[1]

BY

W. MAGNUS

The problem of diffraction of a plane scalar (acoustical) wave by a circular aperture in an infinite plane screen leads to the following boundary-value problem: Find a solution u of $\nabla^2 u + k^2 u = 0$ (with a constant k) with the following properties: If z,ρ,θ are cylindrical coordinates, u satisfies the differential equation everywhere except for $z = 0$ and $\rho \geqq a > 0$. For $z = 0$, $\rho > a$, u vanishes, and for $z = 0$, $0 \leqq \rho < a$, $\partial u/\partial z$ assumes a given constant value $v_0 \neq 0$. At infinity, u satisfies a Sommerfeld-radiation condition. Then $u = u(\rho,z,\theta)$ is uniquely determined and represents the diffracted field.

Because of the symmetry of the boundary conditions, u does not depend on θ. In the diffracting circular aperture, *i.e.*, for $z = 0$, $0 \leqq \rho \leqq a$, $u = \Phi(\rho)$, where Φ depends on ρ only. It has been shown by Levine and Schwinger [1] that u can be determined completely if $\Phi(\rho)$ or even if $C_0\Phi(\rho)$ is known where C_0 denotes an undetermined constant factor different from zero.

Expanding $\Phi(\rho)$ in a series

$$\Phi(\rho) = -\frac{1}{2} a C_0 \sum_{m=0}^{\infty} x_m \left(1 - \frac{\rho^2}{a^2}\right)^{m+\frac{1}{2}} \tag{1}$$

and applying their "variational method," Levine and Schwinger show that the x_m's satisfy a system of inhomogeneous linear equations

$$\sum_{m=0}^{\infty} l_{n,m}(\beta) x_m = \eta_n, \qquad (n = 0,1,2, \cdots), \tag{2}$$

where the $l_{n,m}$'s are explicitly known power series in $\beta = ka/2$ and where $\eta_n = 1/(n + \frac{3}{2})$. Since the quantity

$$T^* = \sum_{m=0}^{\infty} \frac{x_m}{m + \frac{3}{2}} \tag{3}$$

determines the "transmission coefficient" of the diffracted wave which is used in deriving (2), only those solutions of (2) will be considered for which the series in (3) converges. Then the series in (1) converges for $0 < \rho < a$; since $\Phi(\rho)$ must be finite and continuous for $\rho = 0$, we also assume that

[1] Abstract. This work was performed at the Washington Square College of Arts and Science, New York University, and was supported in part by Contract AF-19(122)-42 with the U.S. Air Force through the sponsorship of the Geophysics Research Division, Air Force Cambridge Research Center, Air Matériel Command.

$$\lim_{\epsilon\to 0} \sum_{m=0}^{\infty} x_m(1-\epsilon)^m = C_1 \tag{4}$$

exists.

The $l_{n,m}$'s are given by

$$l_{n,m}(\beta) = \sum_{q=0}^{\infty} l_{n,m}^{(q)}\beta^q,$$

where for $q = 2p$, $p = 0, 1, 2, \cdots$,

$$l_{n,m}^{(2p)} = \frac{(-1)^p\pi^{\frac{1}{2}}}{4p!} \frac{\Gamma(m+\frac{3}{2})\Gamma(n+\frac{3}{2})\Gamma(m+n+2p+1)}{\Gamma(m+p+1)\Gamma(n+p+1)\Gamma(m+n+p+\frac{5}{2})},$$

and for $q = 2p + 3$,

$$l_{n,m}^{(2p+3)} = \frac{i(-1)^p\pi^{\frac{1}{2}}}{4\Gamma(p+\frac{5}{2})} \frac{\Gamma(m+\frac{3}{2})\Gamma(n+\frac{3}{2})\Gamma(m+n+2p+4)}{\Gamma(m+p+\frac{5}{2})\Gamma(n+p+\frac{5}{2})\Gamma(m+n+p+4)}.$$

For $p = 1$, $l_{n,m}^{(1)} = 0$. Denoting by $L^{(p)}$ the matrix with the element $l_{n,m}^{(p)}$ in its $(n+1)$st row and $(m+1)$st column ($n,m = 0,1,2, \cdots$), it can be shown that for $p = 0, 1, 2, 3, \cdots$

$$L^{(p)} = L^{(0)}S^{(p)},$$

where $S^{(0)}$ is the identity and where the matrices $S^{(p)}$ are known explicitly. For odd values of p, all the elements in $S^{(p)}$ except those in the first $(p-1)/2$ rows are zero, and for even values of p, the element in the $(1+n)$th row and $(1+m)$th column of $S^{(p)}$ is zero if $2n < p + 2m$. The $S^{(p)}$'s are bounded matrices, and estimates for their upper bounds can be found. From this it follows:

If the linear equations

$$\sum_{m=0}^{\infty} l_{n,m}^{(0)}x_n^{(0)} = \eta_n \tag{5}$$

have a solution $x_m^{(0)}$ satisfying the conditions implied by (3) and (4), then the x_m's can be represented by power series in β,

$$x_m = \sum_{p=0}^{\infty} x_m^{(p)}\beta^p,$$

which converge at least in a certain range $0 \leqq \beta \leqq \beta_0$. The $x_m^{(p)}$'s can be derived from the $x_m^{(0)}$'s by finite-recurrence relations. In the particular case where $\eta_m = 1/(m+\frac{3}{2})$ we have $x_0^{(0)} = 8/\pi$, $x_m^{(0)} = 0$ for $m > 0$, and at most the first $p + 1$ of the numbers $x_m^{(p)}$ are different from zero. They can be determined by using only the first $p + 1$ of equations (2) and neglecting all terms involving x_{p+1}, x_{p+2}, $\cdots$ and all terms involving β^{p+1}, β^{p+2}, $\cdots$. The x_m's satisfy the conditions implied by (3) and (4).

The factorization of the $L^{(p)}$'s into a product $L^{(0)}S^{(p)}$ shows that the solution of (2) depends largely on the inversion of the matrix $L^{(0)}$, which represents the static case $k = 0$ (or $\beta = 0$). For the limiting case $\beta \to \infty$, Levine and Schwinger have shown that the variational method leads to a system of linear equations

$$\sum_{m=0}^{\infty} l_{n,m}^{(\infty)} x_m^{(\infty)} = C_2 \eta_n, \tag{6}$$

where C_2 is a constant and where

$$l_{n,m}^{(\infty)1} = (n + m + 2)^{-1}. \tag{7}$$

It can be shown that both (6) and (5) are equivalent to the solution of a problem of moments. Consequently, it may happen that (5) or (6) does not have any solutions and in particular does not have a solution satisfying the conditions implied by (3) and (4). But even if (5) or (6) has a solution for a particular choice of the η_m's, no solution will exist if only a finite number of these η_m's is changed. As a matter of fact, (6) does not have any solution at all if

$$\eta_n = \frac{C_2}{n + \frac{3}{2}}.$$

Nevertheless, the quantity T^* can still be determined and even computed from (5), and since it is the aim of the variational method to calculate the transmission coefficient (that is, T^*), the method still works in this case. To prove this, (5) and (7) are shown to arise from certain integral equations.

If we define

$$f(v) = \sum_{m=0}^{\infty} x_m^{(\infty)} v^m, \qquad h(z) = C_2 \sum_{n=0}^{\infty} \eta_n z^n,$$

then the integral equation

$$\int_0^1 f(v) v (1 - vz)^{-1} \, dv = h(z) \tag{8}$$

has a solution $f(v)$ which is analytic for $0 \leqq |v| < 1$ if (6) has a solution for which T^* and C_1 in (3) and (4) exist. But (8) may have a solution which is not analytic for $|v| < 1$ and which therefore does not lead to a solution of (6). This is due to the fact that the integral operator in (8) is an extension of the matrix operator in (6).

In particular, let

$$\eta_n = \frac{1}{n + \frac{3}{2}}$$

or

$$h(z) = C_2 \int_0^1 v^{\frac{1}{2}} (1 - vz)^{-1} \, dv.$$

In this case, $f(v) = C_2 v^{-\frac{1}{2}}$, and although there do not exist any $x_n^{(\infty)}$'s satisfying (6), the quantity T^* can still be defined by writing

$$\sum_{m=0}^{\infty} \frac{x_m^{(\infty)}}{m + \frac{3}{2}} = \int_0^1 f(v)v^{\frac{1}{2}}\,dv = T^*. \tag{9}$$

This gives $T^* = C_2$. Actually T^* can be computed approximately by solving a finite number of the equations in (6) for an equal number of unknowns and by introducing the "approximate" values of the $x_n^{(\infty)}$'s thus obtained into the left-hand side of (9).

Bibliography

1. H. Levine and J. Schwinger, *On the theory of diffraction by an aperture in an infinite plane screen*, Phys. Rev. vol. 74 no. 8 (October, 1948) pp. 958–974.

New York University,
New York, N.Y.

ON THE COUPLING OF TWO HALF PLANES[1]

BY

ALBERT E. HEINS AND HERMAN FESHBACH

1. Introduction. This paper is a continuation of some earlier work which dealt with the coupling of two acoustical ducts of rectangular cross section whose walls were composed of different acoustical materials [1]. We shall be concerned here with the effect of a plane wave incident upon two semi-infinite half planes joined along a line $x = 0$. Each of these half planes is assumed to have an acoustical property which may be described by a complex parameter, the admittance of the material. An approximate solution to this problem has been given by Morse and Bolt [2] for the special case of normal incidence;

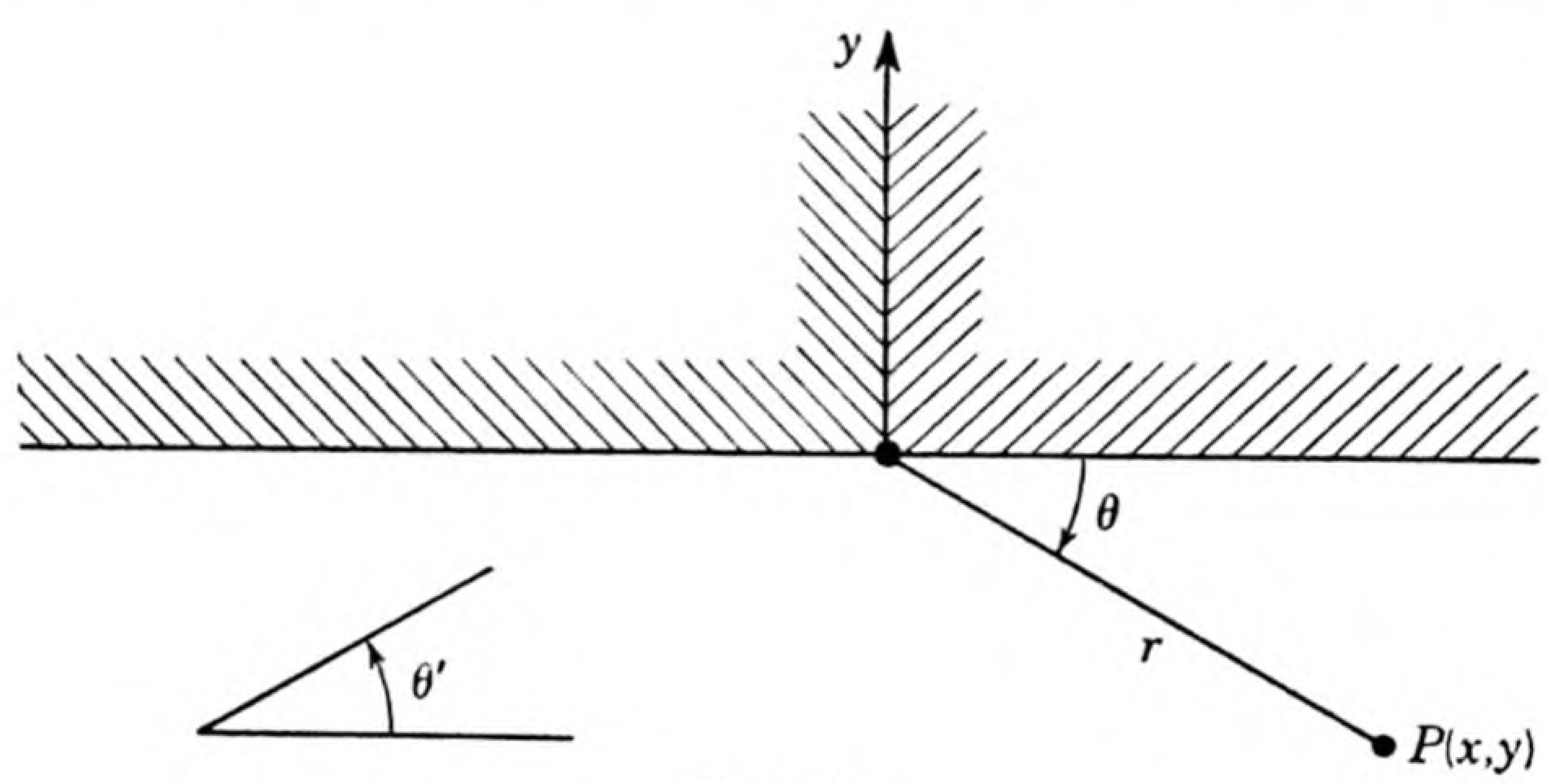


FIG. 1

however, these workers did not find the scattered field, nor did they indicate the nature of the acoustic potential along the two half planes. We shall solve this problem by an appeal to the theory of integral equations. This formulation has been indicated in [2], where Y_1 and Y_2, the admittance parameters, are considered functions of a coordinate. We shall consider the admittance parameters as constants, and hence the integral equation is one of the Wiener-Hopf type [3]. The complex Fourier transform is the appropriate tool with which to solve this problem, and we shall see that our methods lean heavily on the theory of functions of a complex variable.

2. The Formulation of the Problem. Let us consider the half space $y \leq 0$ whose boundary $y = 0$ is made of two different acoustical materials. A plane wave whose propagation normal makes an angle θ' with respect to the positive x-axis is incident upon this structure. (See Fig. 1.) The steady-state velocity potential $\phi(x,y)$, with the time dependence suppressed, satisfies the two-dimensional wave equation [2].

$$\phi_{xx} + \phi_{yy} + k^2\phi = 0. \tag{1}$$

[1] This paper is based on research conducted in part under Contract No. DA-36-061-Ord-113 between the Office of Ordnance Research of the Department of the Army and Carnegie Institute of Technology.

We shall assume for the present that the half space is slightly absorbing and that k therefore possesses a small positive imaginary part.

The acoustical surfaces $y = 0, x \gtrless 0$ may be described in terms of a complex parameter Y, the admittance of the surface. We shall denote the admittance of $x < 0, y = 0$ by Y_1 and that of $x > 0, y = 0$ by Y_2. Y may be written in terms of its real and imaginary parts as $Y = \gamma - i\sigma$. γ is the conductance and σ is the susceptance of the material. γ is nonnegative, and if it does not vanish, absorption results. For such acoustical surfaces as we have considered, we shall employ the boundary conditions [2].

$$\phi_y = i\omega\rho Y_1\phi, \qquad (y = 0, x < 0), \tag{2}$$

and

$$\phi_y = i\omega\rho Y_2\phi, \qquad (y = 0, x > 0) \tag{3}$$

Here ρ is the density of the medium $y \leq 0$, and ω is the angular frequency of the time variation.

It is well known that, for a half space $y < 0$, the scattered part of $\phi(x,y)$, which we denote here by $\phi^s(x,y)$, has the following representation:

$$\phi^s(x,y) = \int_{-\infty}^{\infty} [G(x,y,x',0)\phi^s_{y'}(x,0) - \phi^s(x',0)G_{y'}(x,y,x',0)]\,dx', \tag{4}$$

where G is a Green's function associated with equation (1). Since there are no contributions from infinity, this has been derived subject to the conditions that ϕ^s as well as G satisfy the re-radiation condition of Sommerfeld for $r = (x^2 + y^2)^{\frac{1}{2}} \to \infty$. Now $\phi = \phi^s + \phi^i$, where ϕ^i is the incident plane-wave field so that equation (4) becomes

$$\begin{aligned}\phi^s(x,y) = \int_{-\infty}^{\infty} \{G(x,y,x',0)[\phi_{y'}(x',0) - ik_y\phi^i(x',0)]& \\ - [\phi(x',0) - \phi^i(x',0)]G_{y'}(x,y,x',0)\}\,dx'&.\end{aligned} \tag{5}$$

In particular if we select a Green's function which satisfies the boundary condition $G_y = i\omega\rho Y_1 G$ at $y = 0$, equation (5) reduces to

$$\phi^s(x,y) = \int_0^{\infty} (\beta_2 - \beta_1)G(x,y,x',0)\phi(x',0)\,dx' + A_1\phi^r(x,y).$$

Here $\phi^i = \exp[i(xk_x + yk_y)]$, $\phi^r = \exp[i(xk_x - yk_y)]$, $\beta_1 = i\omega\rho Y_1$, $\beta_2 = i\omega\rho Y_2$, and $A_1 = (k_y - \omega\rho Y_1)/(k_y + \omega\rho Y_1)$. Observe that $\phi^r(x,y)$ is the geometrically reflected plane wave. In particular, the total velocity potential has the representation

$$\phi(x,y) = \int_0^{\infty} (\beta_2 - \beta_1)G(x,y,x',0)\phi(x',0)\,dx' + \phi^i(x,y) + A_1\phi^r(x,y).$$

For $y = 0$, we have the integral equation

$$\phi(x,0) = \int_0^{\infty} (\beta_2 - \beta_1)G(x,0,x',0)\phi(x',0)\,dx' + \phi^i(x,0) + A_1\phi^r(x,0). \tag{6}$$

3. The Green's Function. We shall now give an explicit form for the Green's function which we employed in Sec. 2. Let us recall that $G(x,y,x',y')$ satisfies the nonhomogeneous wave equation

$$G_{xx} + G_{yy} + k^2G = -\delta(x - x')\delta(y - y'), \tag{7}$$

where $\delta(x - x')$ is the Dirac delta function. More precisely, G satisfies the homogeneous wave equation save at the point $x = x'$ and $y = y'$. At that point, G possesses a logarithmic singularity and hence accounts for the representation in equation (4). In addition to satisfying equation (7) we require two further conditions which we have employed in Sec. 2. First we require that G satisfy the boundary condition $G_y = i\omega\rho Y_1 G$, and second we require that for $r = (x^2 + y^2)^{\frac{1}{2}} \to \infty$ we have that G is asymptotic to $\exp(ikr)/r^{\frac{1}{2}}$ in so far as its r variation is concerned. That is, G acts as a two-dimensional source for outward going radiation.

We proceed in a formal manner with this task and indicate how one may verify that all the conditions imposed upon G may be satisfied. A direct application of the complex Fourier transform to equation (7) (for the variable x) gives us

$$\begin{aligned} G(x,y,x',y') &= \frac{1}{2\pi}\int_C \exp\,[iz(x - x') - i\alpha y](A\cos\alpha y' + B\sin\alpha y')\,dz, & (y < y'), \\ &= \frac{1}{2\pi}\int_C \exp\,[iz(x - x') - i\alpha y'](A\cos\alpha y + B\sin\alpha y)\,dz, & (y > y'), \end{aligned} \tag{8}$$

where

$$A = \frac{i}{\alpha + \omega\rho Y_1}, \qquad B = -\frac{\omega\rho Y_1}{\alpha(\alpha + \omega\rho Y_1)}$$

and $\alpha = (k^2 - z^2)^{\frac{1}{2}}$. Here we choose the branch of α which reduces to k for $z = 0$. It remains to describe the nature of the path C.

The representation (8) has been derived subject to the fact that G has a specific asymptotic form for $|x| \to \infty$. This implies that the integrand in (8) is regular in the strip $|\mathrm{Im}\, z| < |\mathrm{Im}\, k|$ provided the denominators in A and B do not vanish in this strip. Subject to this condition, C is therefore a path in this strip. In order to cast (8) into a more useful form, certain deformations of the path C are in order. Before proceeding with this task, however, let us note that (8) may be rewritten as

$$\begin{aligned} G(x,y,x',y') &= \frac{i}{2\pi}\int_C \exp\,[iz(x - x') - i\alpha y]\left[\frac{\exp\,(i\alpha y')}{2\alpha} + \frac{\exp\,(-i\alpha y')}{2\alpha} \right. \\ &\qquad \left. - \frac{\omega\rho Y_1 \exp\,(-i\alpha y')}{\alpha(\alpha + \omega\rho Y_1)}\right] dz, \qquad (y < y'); \\ &= \frac{i}{2\pi}\int_C \exp\,[iz(x - x') - i\alpha y']\left[\frac{\exp\,(i\alpha y)}{2\alpha} + \frac{\exp\,(-i\alpha y)}{2\alpha} \right. \\ &\qquad \left. - \frac{\omega\rho Y_1 \exp\,(-i\alpha y)}{\alpha(\alpha + \omega\rho Y_1)}\right] dz, \qquad (y > y'); \end{aligned} \tag{9}$$

$$= \frac{i}{4} H_0^{(1)}\{k[(x-x')^2 + (y-y')^2]^{\frac{1}{2}}\}$$
$$+ \frac{i}{4} H_0^{(1)}\{k[(x-x')^2 + (y+y')^2]^{\frac{1}{2}}\}$$
$$- \frac{i\omega\rho Y_1}{2\pi} \int_C \frac{\exp\,[iz(x-x') - i\alpha(y+y')]\,dz}{\alpha(\alpha + \omega\rho Y_1)}.$$

The $H_0^{(1)}$ is the Hankel function of the first kind. Incidentally the first Hankel function in the Green's function accounts for the two-dimensional source in the half space $y < 0$, while the other Hankel function and the integral will be shown to combine with the first Hankel function so that the boundary conditions on the line $y = 0$ are satisfied.

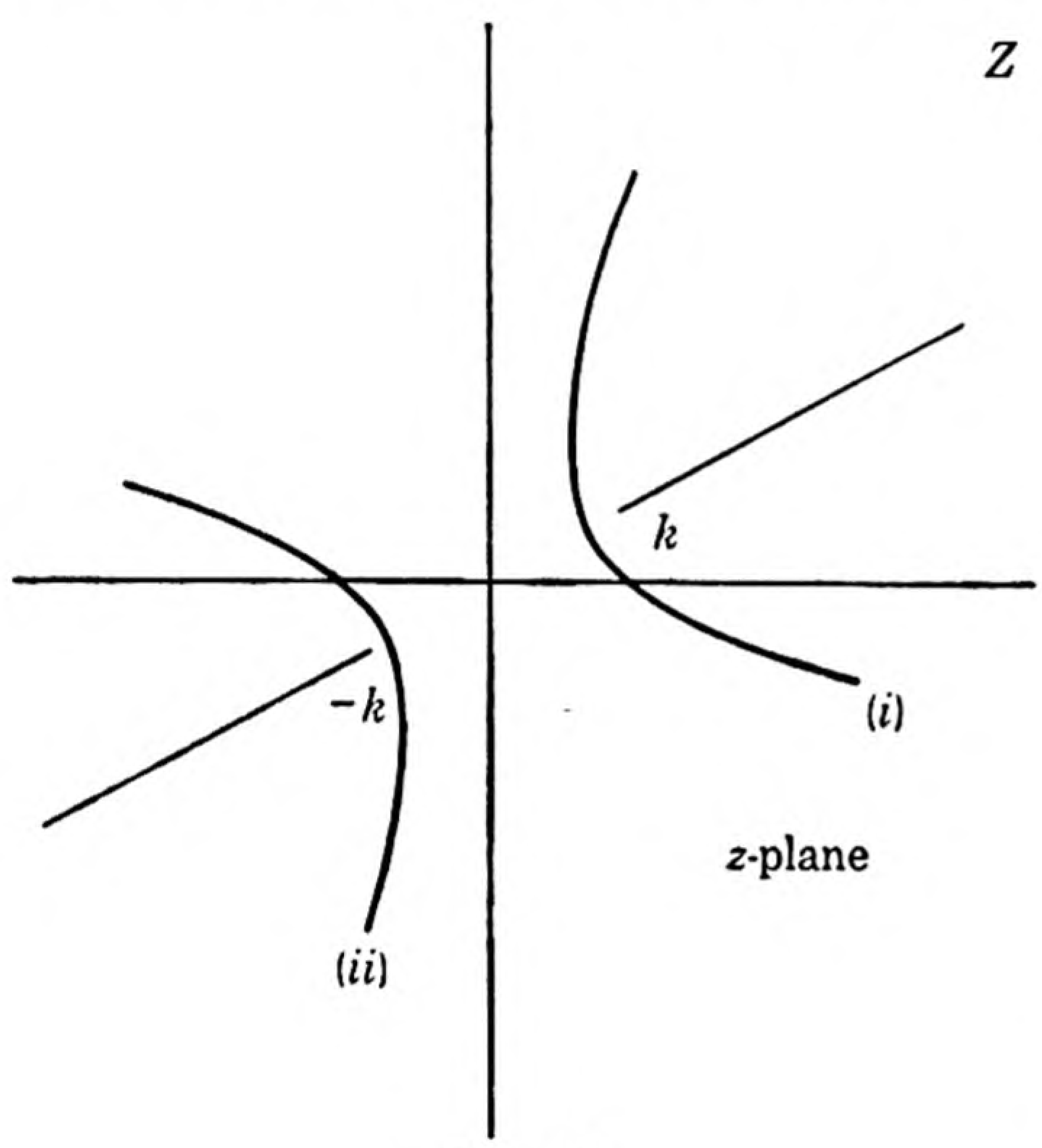


FIG. 2

Deformation of the path C is now required if we are to reduce the integral to a usable form. In order to do this, we introduce radial cuts in the z-plane (see Fig. 2). By Cauchy's theorem, the path C may be deformed into the hyperbolic path [4], $z = k \cos(\psi + i\tau)$, $-\infty < \tau < \infty$. The angle ψ is defined by

$$x - x' = R \cos \psi,$$
$$y + y' = R \sin \psi.$$

The branch of α which we desire is equal to $-k \sin(\psi + i\tau)$, and hence there are no critical points in the denominator of the integrand. More precisely, the denominator can vanish only if

$$k_1 \sin\psi \cosh\tau - k_2 \cos\psi \sinh\tau = \omega\rho\gamma_1$$
$$k_2 \sin\psi \cosh\tau + k_1 \cos\psi \sinh\tau = \omega\rho\sigma_1$$

In these equations we have a measure of k_2 small. If $k_2 \cos\psi < -k_1 \sin\psi$, then there is no real value of τ which satisfies the first equation since γ is positive and the left side will always be negative in the interval $-\pi < \psi < 0$. Hence the integral along C is the negative of the integral along the hyperbolic arc since the contributions from the closing circular are vanished for $x > x'$ when hyperbola (i) is employed. A similar situation holds for $x < x'$ when hyperbola (ii) is used. It follows that

$$(10) \quad \frac{1}{2\pi i} \int_C \frac{\exp\,[iz(x-x') - i\alpha(y+y')]\,dz}{\alpha(\alpha + \omega\rho Y_1)}$$
$$= -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp\,(ikR\cosh\tau)\,d\tau}{\omega\rho Y_1 - k\sin(\psi + i\tau)}$$
$$= -\frac{1}{\pi} \int_0^{\infty} \frac{[\exp\,(ikR\cosh\tau)](\omega\rho Y_1 - k\sin\psi\cosh\tau)\,d\tau}{(\omega\rho Y_1)^2 - 2k\omega\rho Y_1 \sin\psi\cosh\tau + k^2(\cosh^2\tau - \cos^2\psi)}.$$

Because of the presence of the small positive imaginary part in k, the integral in (10) is absolutely convergent for $R \neq 0$. It is therefore permissible to differentiate under the integral sign. Let us denote this integral by $f(\psi,R)$. Then $f_y - i\omega\rho Y_1 f$ is equal to

$$\frac{i}{\pi}\int_0^\infty \frac{[\exp(ikR\cosh\tau)](\omega\rho Y_1 - k\sin\psi\cosh\tau)^2\,d\tau}{(\omega\rho Y_1)^2 - 2k\omega\rho Y_1\sin\psi\cosh\tau + k^2(\cosh^2\tau - \cos^2\psi)}$$
$$-\frac{i}{\pi}\int_0^\infty \exp(ikR\cosh\tau)\psi_y\frac{\partial}{\partial\psi}\left[\frac{\omega\rho Y_1 - k\sin\psi\cosh\tau}{(\omega\rho Y_1)^2 - 2k\omega\rho Y_1\sin\psi\cosh\tau + k^2(\cosh^2\tau - \cos^2\psi)}\right]d\tau$$

provided $R \neq 0$. This in turn may be rewritten as

$$\frac{i}{\pi}\int_0^\infty \exp(ikR\cosh\tau)\,d\tau$$
$$-\frac{ik^2\cos^2\psi}{\pi}\int_0^\infty \frac{\sinh^2\tau[\exp(ikR\cosh\tau)]\,d\tau}{(\omega\rho Y_1)^2 - 2k\omega\rho Y_1\sin\psi\cosh\tau + k^2(\cosh^2\tau - \cos^2\psi)}$$
$$-\frac{1}{\pi}\int_0^\infty \exp(ikR\cosh\tau)\psi_y\frac{\partial}{\partial\psi}\left[\frac{\omega\rho Y_1 - k\sin\psi\cos\tau}{(\omega\rho Y_1)^2 - 2k\omega\rho Y_1\sin\psi\cosh\tau + k^2(\cosh^2\tau - \cos^2\psi)}\right]d\tau.$$

Upon integrating the second integrand by parts we find that, after some simplification, only the first term remains. Hence the operation $-i\omega\rho Y_1 + \partial/\partial y$ applied to the Green's function vanishes when evaluated at either y or $y' = 0$. We have thus shown that the boundary condition at $y = 0$ is satisfied. In a similar fashion we can show that the Green's function satisfies the homogeneous two-dimensional wave equation save at the points $x = x'$, $y = y'$. Indeed for $y < 0$, $y' < 0$, the last two terms in the Green's function (9) satisfy the homogeneous wave equation, while the first term accounts for the unit source.

One condition remains to be discussed—*i.e.*, the condition of outward-going radiation. Again for $r = (x^2 + y^2)^{\frac{1}{2}} \to \infty$ the first two terms of the Green's function (9) obey the requirement that they be asymptotic to $\exp(ikr)/r^{\frac{1}{2}}$ with the point $P'(x',y')$ located in the finite part of the plane. The third term requires some further comment. Let us write $\cosh\tau = 1 + v$. Then the integral in the Green's function transforms into

(11)
$$-\frac{1}{\pi}\int_0^\infty \frac{\exp(ikRv + ikR)[\omega\rho Y_1 - k(1+v)\sin\psi]v^{-\frac{1}{2}}\,dv}{(2+v)^{\frac{1}{2}}[k^2(1+v)^2 - 2k\omega\rho Y_1\sin\psi(1+v) + (\omega\rho Y_1)^2 - k^2\cos^2\psi]}.$$

In this form it is clear that (11) approaches zero as $|kR| \to \infty$ for $\operatorname{Im} k \geq 0$ as a direct consequence of the Riemann-Lebesgue lemma. But more than this can be said, for it is possible to give the precise asymptotic form of the integral as $|kR| \to \infty$, $\operatorname{Im} k \geq 0$. There is another lemma, due to G. N. Watson [5],

which states that since the integrand in (11) is regular and of bounded variation for $v \geq 0$ save for the factor $v^{-\frac{1}{2}}$, we have that for $|k|R \to \infty$, $\text{Im } k \geq 0$, (11) is asymptotic to

$$-\frac{1}{\sqrt{2\pi}} \frac{\exp(ikR)\exp(i\pi/4)}{\sqrt{kR}\,(\omega\rho Y_1 - k \sin\psi)},$$

which is clearly of the required form when the point $P'(x',y')$ is located in the finite part of the plane. We have thus verified formal derivation of the Green's function.

4. The Fourier Transform of Equation (6). In order to cast equation (6) into a form which is amenable to the methods of Wiener and Hopf, we write it as follows,

$$\phi_0(x) + \phi_1(x) = \phi_2(x) + (\beta_2 - \beta_1)\int_{-\infty}^{\infty} \phi_0(x')G(x,0,x',0)\,dx, \tag{12}$$

where now

$$\begin{aligned} \phi_0(x) &\equiv \phi(x,0), && (x > 0), \\ &\equiv 0, && (x < 0), \\ \phi_1(x) &\equiv \phi(x,0), && (x < 0), \\ &\equiv 0, && (x > 0), \\ \phi_2(x) &\equiv (1 + A_1)\exp(ik_x x), && (x > 0), \\ &\equiv 0, && (x < 0). \end{aligned}$$

Clearly then $\phi_1(x)$ is of the order

$$\int_0^{\infty} \frac{\exp[ik(x' - x)]\phi_0(x')\,dx'}{(x' - x)^{\frac{1}{2}}}$$

as $x \to -\infty$. It follows then that $\phi_1(x)$ is dominated by the exponential growth, $\exp(k_2 x)$, where $k_1 + ik_2 = k$. Furthermore, the above integral will converge absolutely if $\phi_0(x) = O[\exp(-i\,\delta x)]$, $x \to \infty$, and $\text{Im } \delta < k_2$. We shall assume for the present that $\phi_0(x)$ and $\phi_1(x)$ are integrable for finite x and verify this assumption with the solution of equation (12). With this integrability assumption, we can say something about the regions of regularity of the unilateral Fourier transforms of $\phi_1(x)$ and $\phi_0(x)$. Since the exponential growth of a function is the determining factor in the location of the half planes of regularity (assuming integrability for finite x), we have that

$$\Phi_0(z) = \int_0^{\infty} \exp(-izx)\phi_0(x)\,dx$$

is regular in the lower half plane $\text{Im } z < \text{Im } \delta$, while

$$\Phi_1(z) = \int_{-\infty}^{0} \exp(-izx)\phi_1(x)\,dx$$

is regular in the upper half plane $\text{Im } z > -k_2$.

The function $G(x,0,0,0)$ has the bilateral transform $i/(\omega\rho Y_1 + \alpha)$, where the root of the radical α has been chosen to reduce to k for $z = 0$. In view of the fact that this transform has only branch points as singularities, it is regular in the strip $-\text{Im } k < \text{Im } z < \text{Im } k$. Indeed it is actually regular in the cut plane, with such cuts as the radial ones described in Sec. 3 as possible useful ones. Finally $\phi_2(x)$ has the unilateral transform $(1 + A_1)/i(z - k_x)$, which is regular in the lower half plane $\text{Im } z < k_2 \cos \theta'$. There is then a common strip of analyticity for these transforms that is

$$-k_2 < \text{Im } z < k_2 \cos \theta' < \text{or Im } \delta.$$

The right side of this inequality is to be interpreted as the smaller of the two terms. Because of the existence of this common strip, we may apply the Fourier transform theorem to equation (12) to obtain

$$\Phi_0(z) + \Phi_1(z) = \frac{1 + A_1}{i(z - k_x)} + \frac{i(\beta_2 - \beta_1)\Phi_0(z)}{\alpha + \omega\rho Y_1}$$

or

$$\Phi_1(z) + \Phi_0(z)\,\frac{\alpha + \omega\rho Y_2}{\alpha + \omega\rho Y_1} = \frac{1 + A_1}{i(z - k_x)}. \tag{13}$$

The explicit solution of equation (12) depends on our ability to decompose equation (13) into two parts, one of which is regular in the upper half plane $\text{Im } z > -k_2$, while the other is regular in the lower half plane $\text{Im } z < \text{Im } \delta$ or $\text{Im } k_x$. Our next task is therefore the decomposition of

$$K(z) = \frac{\alpha + \omega\rho Y_2}{\alpha + \omega\rho Y_1}$$

into two such functions as we have just described.

5. The Decomposition of $K(z)$. The required factoring of $K(z)$ may be carried out with the aid of Cauchy's second integral theorem. In order to accomplish this, we note that the multiplicative decomposition of $K(z)$ is essentially equivalent to the additive decomposition of the logarithm of $K(z)$. From the calculational point of view, however, it is better to examine the additive decomposition of the derivative of the logarithm of $K(z)$. Let us note first that $K(z)$ has two critical points at $\pm k$. We shall consider here, however,

$$L(z) = 1 + \frac{\omega\rho Y_1}{(k^2 - z^2)^{\frac{1}{2}}}$$

and

$$M(z) = 1 + \frac{\omega\rho Y_2}{(k^2 - z^2)^{\frac{1}{2}}}$$

individually. After having obtained the required decomposition of each of these functions, we may form their ratio to find the decomposition of $K(z)$.

Within the strip indicated in Fig. (3), $L(z)$ as well as its logarithm is regular. The arguments of the radicals on the branch cuts (which are now drawn parallel to the horizontal axis of the z-plane) are indicated in the figure. Then

$$\frac{L'(z)}{L(z)} = \frac{1}{2\pi i}\int_C \frac{\omega\rho Y_1 t\,dt}{[(k^2-t^2)^{\frac{1}{2}}+\omega\rho Y_1](k^2-t^2)(t-z)}$$

$$= \frac{\omega\rho Y_1 z}{[(k^2-z^2)^{\frac{1}{2}}+\omega\rho Y_1](k^2-z^2)}.$$

The path C is a rectangular one which includes the point z and is indented above the branch point at $z = -k$ and below the branch point at $z = k$. The vertical

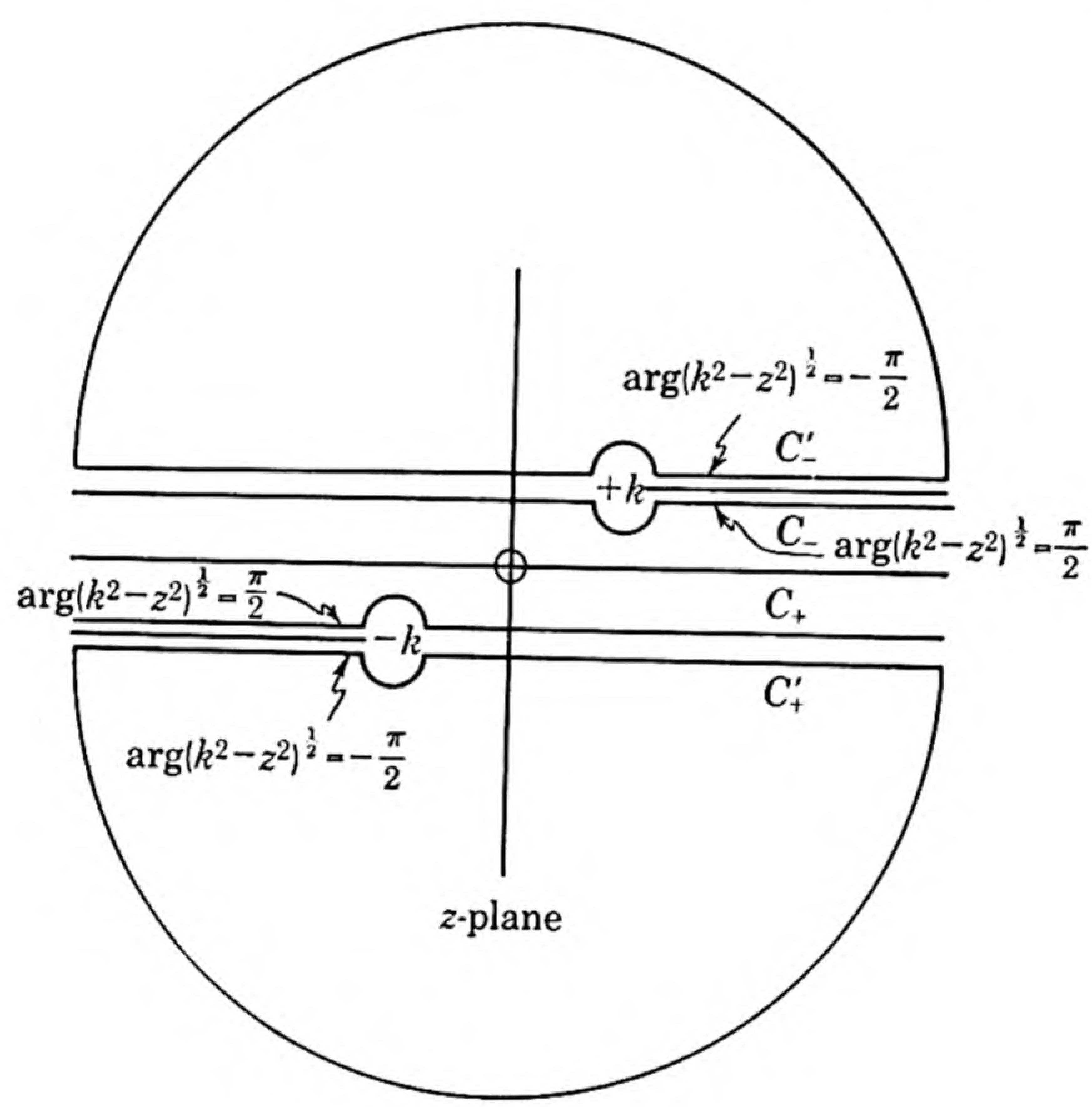


FIG. 3

sides of this rectangle give rise to terms which vanish as they recede to infinity in either direction, in view of the fact that the integrand is $0(t^{-3})$, $|t| \to \infty$. Clearly, then, the integral along the lower side of the rectangle, for z not on C but inside the strip, provides us with a function regular in the upper half plane $\operatorname{Im} z > -k_2$, while the integral along the upper side of the rectangle for z not on C provides us with a function regular in the lower half plane $\operatorname{Im} z < k_2$. Let us denote these component paths of C by C_+ and C_-, respectively. Useful forms for the integrals along C_+ and C_- may be obtained by appropriate deformations.

We consider the integral along C_+ first. Let us introduce a closed path C'_+ which passes under the left branch cut, is indented below the branch cut $t = -k$, and is closed by a semicircular arc in the lower half plane. The semi-circular portion of C'_+ provides a vanishing contribution in the limit as its radius becomes infinite. Hence the horizontal portion of C'_+ combines with

C_+ in the following manner when we take into account that we have a simple pole as well as a branch point at $t = -k$:

$$\begin{aligned}\frac{L'_+(z)}{L_+(z)} &= \frac{1}{2\pi i}\int_{C_+} = \frac{1}{2\pi i}\int_{C_+} + \frac{1}{2\pi i}\int_{C'_+} \\ &= \frac{1}{\pi}\int_{-\infty}^{-k} \frac{\omega\rho Y_1 t\, dt}{(t^2 + \omega^2\rho^2 Y_1^2 - k^2)(t^2 - k^2)^{\frac{1}{2}}(t - z)} - \frac{1}{2(z + k)} \\ &= -\frac{1}{2(z + k)} - \frac{\omega\rho Y_1}{\pi}\left[\frac{\psi(z) - \psi(\alpha_1)}{z - \alpha_1} + \frac{\psi(z) - \psi(-\alpha_1)}{z + \alpha_1}\right].\end{aligned}$$

Here $\psi(z)$ is given by the following expression,

$$\psi(z) = (k^2 - z^2)^{-\frac{1}{2}} \arctan\left[(k - z)^{\frac{1}{2}}(k + z)^{-\frac{1}{2}}\right],$$

while $L_+(z)$ is that multiplicative factor of $L(z)$ which is regular in the upper half plane $\operatorname{Im} z > -k_2$ and $L'_+(z)$ is its derivative. $\alpha_1 = (k^2 - \omega^2\rho^2 Y_1^2)^{\frac{1}{2}}$, and it is the root which has a positive imaginary part. An equivalent form for $\psi(z)$ which is useful for $|z| \to \infty$, $\operatorname{Im} z$ appropriately limited, is

$$\psi(z) = \frac{(z^2 - k^2)^{-\frac{1}{2}}}{2i} \ln \frac{z + (z^2 - k^2)^{\frac{1}{2}}}{k}.$$

For $|z| \to \infty$, $\operatorname{Im} z > -k_2$ we have that

$$\psi(z) = \frac{z^{-1}}{2i} \ln \frac{2z}{k} + 0(z^{-2}).$$

We have finally upon integrating that

$$L_+(z) = (z + k)^{-\frac{1}{2}} \exp\left\{\frac{-\omega\rho Y_1}{\pi}\int^z \left[\frac{\psi(t) - \psi(\alpha_1)}{t - \alpha_1} + \frac{\psi(t) - \psi(-\alpha_1)}{t + \alpha_1}\right] dt\right\}$$

up to an arbitrary multiplicative constant. By a similar calculation in the upper half plane we find that $L_-(z) = L_+(-z)$. The decomposition of $M(z)$ is similar to that of $L(z)$ save for the fact that Y_1 is replaced by Y_2. Hence $K_+(z)$ is simply the ratio of $M_+(z)$ to $L_+(z)$. We shall record here, for future use, the asymptotic forms of $K_+(z)$ and $K_-(z)$ for $z \to \infty$ and $\operatorname{Im} z$ suitably limited. Using the asymptotic forms we gave for $\psi(z)$, it is clear that

$$K_+(z) = c\left[1 + \frac{\omega\rho}{\pi}(Y_2 - Y_1)\frac{\ln z}{z}\right],$$

$|z| \to \infty$, $\operatorname{Im} z > -k_2$, while

$$K_-(z) = c_1\left[1 - \frac{\omega\rho}{\pi}(Y_2 - Y_1)\frac{\ln z}{z}\right],$$

$|z| \to \infty$, $\operatorname{Im} z < k_2$. The constants c_1 and c satisfy the condition that $cc_1 = 1$.

6. The Fourier-transform Solution of Equation (13). Equation (13) may now be written in terms of $K_-(z)$ and $K_+(z)$ as

$$(14)\qquad -\frac{\Phi_1(z)}{K_+(z)}+\frac{1+A_1}{i(z-k_x)}\left[\frac{1}{K_+(z)}-\frac{1}{K_+(k_x)}\right]=K_-(z)\Phi_0(z)-\frac{1+A_1}{i(z-k_x)K_+(k_x)},$$

where the left side is regular in the upper half plane $\text{Im } z > -\text{Im } k$, while the right side is regular in the lower half plane $\text{Im } z < -\text{Im } \delta$ or $\text{Im } k_x$. Since both sides are regular in a common strip, the left side is the analytic continuation of the right side and therefore each side of equation (14) is regular in the entire z-plane, *i.e.*, each side is an entire function $E(z)$. The determination of $E(z)$ in this case is a straightforward task since $K_-(z)$ and $K_+(z)$ are bounded for $|z| \to \infty$, $\text{Im } z$ suitably limited, and since $\Phi_1(z)$ and $\Phi_0(z)$ approach zero for $|z| \to \infty$, $\text{Im } z$ suitably limited. Indeed from these facts we note that $E(z)$ approaches zero for $|z| \to \infty$, and an immediate application of Liouville's theorem on bounded functions tells us that $E(z)$ vanishes identically. This gives us then

$$(15)\qquad \Phi_0(z) = \frac{1+A_1}{i(z-k_x)K_-(z)K_+(k_x)}.$$

From equation (15) we see that $\Phi_0(z) = 0(z^{-1})$ as $|z| \to \infty$, $\text{Im } z < \text{Im } k_x$, and therefore $\phi_0(x) = 0(1)$, $x \to 0^+$. Incidentally $\Phi_0(z)$ is regular in the lower half plane $\text{Im } z < \text{Im } k_x$, since the next singularity of $\Phi_0(z)$ is k. If we examine the expression for $\Phi_1(z)$ for $|z| \to \infty$, $\text{Im } z > -\text{Im } k$, we find that

$$\Phi_1(z) = \frac{1+A_1}{iz} - \frac{1+A_1}{izc_1K_+(k_x)}$$

up to terms of $0(1/z)$. This tells us that

$$\phi_1(x) = -(1+A_1) + \frac{1+A_1}{c_1K_+(k_x)}, \qquad (x \to 0^-).$$

But $\phi_1(x) = \phi(x,0) - (1 + A_1)$, so that for $x \to 0^-$ we have that

$$\phi(x,0) = \frac{1+A_1}{c_1K_+(k_x)}.$$

A similar calculation for $\phi_0(x)$ gives us the same result so that we see that $\phi(x,0)$ is continuous at $x = 0$.

7. The Far Field. The calculation of the far field, which incidentally will provide us with the diffraction pattern, does not offer any major difficulty. Care is required in the neighborhood of the angle $\theta = -\theta'$, and this brings us again to a problem which has occupied workers in diffraction theory since the days of Sommerfeld's [6] initial contributions. It appears that if the complex integral is cast into an appropriate form, the problem of calculating the far field in the neighborhood of $\theta = -\theta'$ is considerably simplified.

We start by considering representations of $\phi(x,y)$ for $x \gtrless 0$. In the first place we have from the convolution theorem the following Fourier-integral representation for $\phi(x,y)$:

$$\phi(x,y) = \phi^i(x,y) + A_1\phi^r(x,y) + \frac{\beta_2 - \beta_1}{2\pi}(1 + A_1)\int_C \frac{K_{+}^{-1}(k_x)\exp\,[izx - iy(k^2 - z^2)^{\frac{1}{2}}]\,dz}{K_-(z)[\omega\rho Y_1 + (k^2 - z^2)^{\frac{1}{2}}](z - k_x)},$$

where the path C is drawn in the strip of regularity $-\text{Im}\,k < \text{Im}\,z < \text{Im}\,k_x$. Useful forms of this integral may be derived by deforming the path C into a hyperbolic path in the same fashion which we indicated in Sec. 3. We write $x = R\cos\theta$, $y = R\sin\theta$, $-\pi < \theta < 0$, $R \geq 0$, while $z = k\cos(\theta + i\tau)$, $-\infty < \tau < \infty$ on the hyperbolic path. The integral representation for $\phi(x,y)$ then becomes

$$\phi(x,y) = \phi^i(x,y) + D\phi^{(r)}(x,y) + \frac{i(\beta_2 - \beta_1)}{2\pi}\int_{-\infty}^{\infty} \frac{\exp\,(ikR\cosh\tau)K_{+}^{-1}(k_x)\sin(\theta + i\tau)\,d\tau}{K_-[k\cos(\theta + i\tau)][\omega\rho Y_1 - k\sin(\theta + i\tau)][\cos(\theta + i\tau) - \cos\theta']},$$

where $D = A_1$ for $-\pi < \theta < -\theta'$ and $D = A_2$ for $-\theta' < \theta < 0$. For $\theta = -\theta'$ the integral is defined in the sense of a Cauchy principal value.

We now turn our attention to the integral

$$\int_{-\infty}^{\infty} \frac{\exp\,(ikR\cosh\tau)K_{+}^{-1}(k_x)\sin(\theta + i\tau)\,d\tau}{K_-[k\cos(\theta + i\tau)][\omega\rho Y_1 - k\sin(\theta + i\tau)][\cos(\theta + i\tau) - \cos\theta']}$$

in the limit that $|kR| \to \infty$. This integral may be rewritten as

$$\int_{-\infty}^{\infty} \frac{\exp\,(ikR\cosh\tau)K_{+}^{-1}(k_x)\sin(\theta - i\tau)\,d\tau}{K_-[k\cos(\theta - i\tau)][\omega\rho Y_1 - k\sin(\theta - i\tau)][\cos(\theta - i\tau) - \cos\theta']}$$

and in turn may be rewritten as the sum of the following two integrals:

$$\int_{-\infty}^{\infty} \frac{[f(\tau) + f(-\tau)]\exp\,(ikR\cosh\tau)\,[\omega\rho Y_1(sc - sc'A) - ks^2A(c - c'A) - kcB^2(A - cc')]\,d\tau}{(A - cc' - ss')(A - cc' + ss')(\omega^2\rho^2Y_1^2 - 2ksA\omega\rho Y_1 + k^2s^2 + k^2B^2)} + \int_{-\infty}^{\infty} \frac{[f(\tau) - f(-\tau)]\exp\,(ikR\cosh\tau)[\omega\rho Y_1(A - cc') + ks(c^2 - A^2)]B\,d\tau}{(A - cc' - ss')(A - cc' + ss')(\omega^2\rho^2Y_1^2 - 2ksA\omega\rho Y_1 + k^2s^2 + k^2B^2)}. \tag{16}$$

The following abbreviations have been employed in the last two integrals:

$$f(\tau) = \frac{K_{+}^{-1}(k_x)}{K_-[k\cos(\theta + i\tau)]}, \qquad A = \cosh\tau,$$

$$B = \sinh\tau, \qquad c = \cos\theta, \qquad c' = \cos\theta', \qquad s = \sin\theta, \qquad s' = \sin\theta'.$$

We observe that each of the integrands in (16) is well behaved save possibly for the factor $1/(A - cc' + ss')$. For $\tau = 0$ and $\theta = -\theta'$ this factor is singu-

lar. (The case $\theta = \theta'$ cannot occur here.) In order to isolate this difficulty, we rewrite (16) as

$$\frac{1}{2}\int_{-\infty}^{\infty} \frac{[f(\tau) + f(-\tau)]g(\tau) - 2g(0)f(0)}{A - cc' + ss'} \exp(ikR\cosh\tau)\,d\tau + \frac{1}{2}\int_{-\infty}^{\infty} \frac{2g(0)f(0)\exp(ikR\cosh\tau)}{A - cc' + ss'}\,d\tau + \frac{1}{2}\int_{-\infty}^{\infty} \frac{[f(\tau) - f(-\tau)]h(\tau)}{A - cc' + ss'} \exp(ikR\cosh\tau)\,d\tau$$

where now

$$g(\tau) = \frac{(sc - sc'A)\omega\rho Y_1 - ks^2A(c - c'A) - kcB^2(A - cc')}{(A - cc' - ss')(\omega^2\rho^2Y_1^2 - 2ksA\omega\rho Y_1 + k^2s^2 + k^2B^2)}$$

and

$$h(\tau) = \frac{[\omega\rho Y_1(A - cc') + ks(c^2 - A^2)]B}{(A - cc' - ss')(\omega^2\rho^2Y_1^2 - 2ksA\omega\rho Y_1 + k^2s^2 + k^2B^2)}.$$

The integrands in the first and third integrals are continuous for $\tau = 0$, $\theta = -\theta'$, and an application of Kelvin's method of stationary phase shows us that for $\theta \neq -\theta'$ each of these integrals is $O(1/kR)$ as $|kR| \to \infty$. For $\theta = -\theta'$ we get a contribution $O(1/\sqrt{kR})$ for $|kR| \to \infty$, but we shall not pursue this calculation further.

This leaves us with the second integral, which may be rewritten as

$$g(0)f(0)\int_{-\infty}^{\infty} \frac{[1 - \cosh(\tau/2)]\exp(ikR\cosh\tau)\,d\tau}{A - cc' + ss'} + g(0)f(0)\int_{-\infty}^{\infty} \frac{\cosh(\tau/2)\exp(ikR\cosh\tau)\,d\tau}{A - cc' + ss'}.$$

The first of these integrals again provides us with a contribution $O(1/kR)$ as $|kR| \to \infty$ for $\theta \neq -\theta'$ and $O(1/\sqrt{kR})$ as $|kR| \to \infty$ for $\theta = -\theta'$. The last integral may be expressed in terms of a Fresnel integral as

$$\frac{-i\sqrt{\pi}\exp(i\pi/4)\exp[ikR\cos(\theta + \theta')]}{|\sin(\theta + \theta')/2|}\int_{|2[\sin^2(\theta+\theta')/2]kR|}^{\infty} e^{iw^2}\,dw.$$

This integral provides us with the basis of determining the behavior of $\phi(x,y)$ as $|kR| \to \infty$. If $\theta \neq -\theta'$, the dominant term is $O[\exp(ikR)/\sqrt{kR}]$ for $|kR| \to \infty$. In the neighborhood of $\theta = -\theta'$, we simply have a bounded contribution. Such a situation also arises in the transition from the illuminated to the shadow region in the excitation of an infinite half-plane wave in electromagnetic or acoustic diffraction and has been discussed by Sommerfeld and others. In summary, we have that $\phi(x,y)$ is asymptotic to

$$\phi(x,y) = \phi^i(x,y) + A_1\phi^r(x,y)$$
$$\cdot\ \frac{\exp(i\pi/4)(\beta_2 - \beta_2)g(0)f(0)\sqrt{\pi}\exp[ikR\cos(\theta + \theta')]}{2\pi|\sin(\theta + \theta')/2|}\int_{|2[\sin^2(\theta+\theta')/2]kR|}^{\infty} e^{iw^2}\,dw$$

for $|kR| \to \infty$, $|\theta| > \theta'$ while it is asymptotic to a similar expression for $|\theta| < \theta'$ with the constant A_1 replaced by A_2. For $\theta = -\theta'$ other contributions arise, as we have already noted.

BIBLIOGRAPHY

1. Albert E. Heins and Herman Feshbach, *The coupling of two acoustical ducts*, J. Math. Phys. vol. 26 (1947) pp. 143–155.

2. Phillip M. Morse and Richard Bolt, *Sound waves in rooms*, Rev. Mod. Phys. vol. 16 (1944).

3. Raymond E. A. C. Paley and Norbert Wiener, *The Fourier transform in the complex domain*, Amer. Math. Soc. Colloquium Publications vol. 19 (1934) Chap. IV.

4. Edmund T. Copson, *On an integral equation arising in the theory of diffraction*, Oxford Quart. Math. vol. 17 (1946) pp. 19–34.

5. Edmund T. Copson, *The asymptotic expansion of a function defined by a definite integral or contour integral*, Admiralty Computing Service, 1946.

6. Arnold Sommerfeld, *Mathematische Theorie der Diffraction*, Math. Ann. vol. 47 (1896) pp. 317–374.

CARNEGIE INSTITUTE OF TECHNOLOGY,
PITTSBURGH, PA.
MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASS.

ON THE DIFFUSION OF TIDES INTO PERMEABLE ROCK

BY

G. F. CARRIER AND W. H. MUNK[1]

1. Introduction. In the irrigation wells of the Hawaiian Islands, water-level fluctuations whose frequency components correspond to those of the ocean tides have been observed by Cox [1]. In this paper we shall formulate the problem associated with this phenomenon, assuming the observed ground-water fluctuations to represent a diffusive transmission of the tidal disturbances through the porous volcanic structure of the island. The agreement with observation is fair. The results obtained here may find application in a method for estimating subsurface permeability.

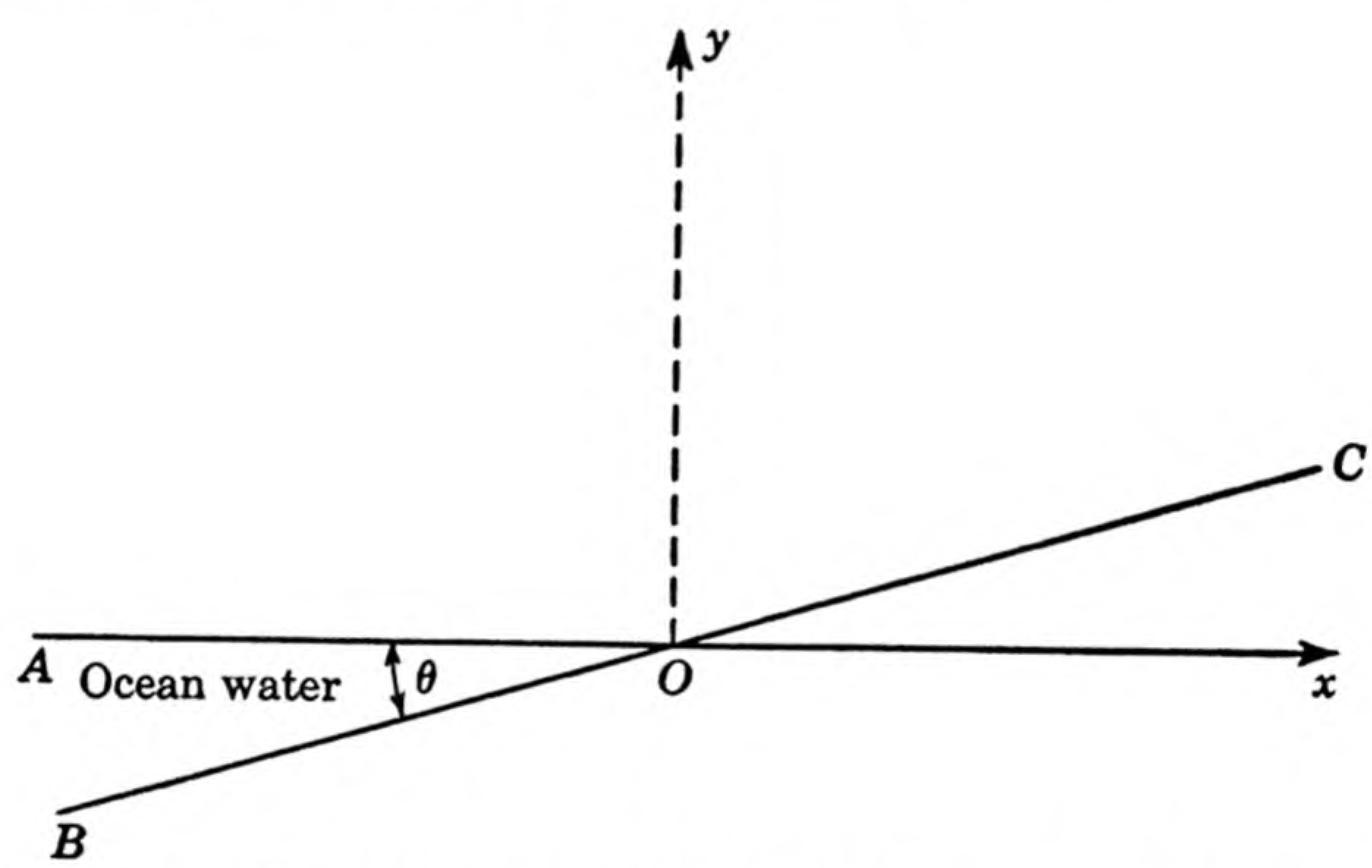


FIG. 1. Below BOC the medium is permeable. The pressure variations on OB depend only on the tidal water-level variations (mean water level: OA).

2. The Formulation of the Problem. Consider the situation in Fig. 1. If the ocean level remained fixed, the pressure distribution in the fluid would be

(1) $$p_0(y) = -\rho_0 g y.$$

However, when the fluid is in motion, the pressure distribution depends on the velocity field. Within the porous medium, we adopt Darcy's law

(2) $$\mathbf{v} = \frac{\mathbf{m}}{\rho_0} = -\frac{k}{\mu}\operatorname{grad}(p - p_0),$$

where p is the pressure, k the permeability of the medium (taken as isotropic and constant), μ the viscosity of the fluid, $\mathbf{m}$ the mass flow per unit area, ρ_0 the mean fluid density, and $\mathbf{v}$ as defined by $\mathbf{m}/\rho_0$ is the "velocity."

Equation (2) replaces the law of conservation of momentum but we must also invoke the mass-conservation law. This takes the form

(3) $$\operatorname{div}\mathbf{m} = -\frac{\partial}{\partial t}A\rho$$

[1] Contribution from the Scripps Institution of Oceanography, new series, 659.

where A is that fraction of the volume which is filled with fluid (*i.e.*, the porosity). We adopt a simple compressibility law for the rock and fluid, *viz.*,

(4) $$\rho A = \rho_0 A_0[1 + \gamma(p - p_0)],$$

where γ is essentially $(\rho_0 c^2)^{-1}$ with c the speed of sound in the fluid. We may now combine the foregoing equations to obtain

(5) $$\Delta q = \frac{\mu\gamma A_0}{k} q_t,$$

where $q = p - p_0$ and Δ is the Laplace operator.

In equation (5) the expression $k/\mu\gamma A_0$ has the dimensions of diffusivity (square centimeters per second), and the equation is of the diffusion type rather than a wave equation. The essential reason for this is that the pressure gradient is proportional to a velocity rather than to an acceleration.

The boundary conditions must now be given. On OB, the sea bottom, we may associate the "pressure" q with the tidal-wave amplitude. In fact, if we decide to consider a single periodic tidal-wave contribution, we write

$$q_{(\text{on } OB)} = q_1 e^{i\omega t}.$$

On the free surface near Ox, we have from equation (2)

(6) $$\eta_t = \frac{\mathbf{j} \cdot \mathbf{v}}{A} = \frac{k}{\mu A} q_y,$$

where $\eta(x,t)$ is the y-coordinate of the free surface. On this surface, however, $p = 0$, and using equation (1),

(7) $$q(x,\eta) = -p_0 = \rho_0 g \eta.$$

We may now use equations (6) and (7) to obtain

(8) $$q_t = -\frac{\rho_0 g k}{\mu A} q_y.$$

In these equations, q should be evaluated at $y = \eta$ but, in the usual linearized way, we apply equation (8) at $y = 0$.

It is now convenient to define some dimensionless quantities, *viz.*:

$$\tau = \omega t, \qquad y\frac{\mu A \omega}{\rho_0 g k} = \frac{y}{L} = y', \qquad \frac{x}{L} = x',$$
$$\epsilon = \frac{\rho_0^2 g^2 k \gamma}{\mu A_0 \omega}.$$

We then have

(9) $$q_{x'x'} + q_{y'y'} = \epsilon q_\tau,$$
(10) $$q(OB) = e^{i\tau},$$
(11) $$q_\tau(x',0,\tau) + q_{y'}(x',0,\tau) = 0.$$

If we now omit the prime on x and y and consider the periodic response of the system, we may write

$$q(x,y,t) = \varphi(x,y)e^{i\tau}$$

and equations (9) to (11) become

$$\Delta\varphi - i\epsilon\varphi = 0, \tag{12}$$

$$\varphi = 1 \text{ on } OB, \tag{13}$$

$$\varphi_y(x,0) + i\varphi(x,0) = 0, \qquad (x > 0). \tag{14}$$

This problem will be solved for two orientations of the line OB: (1) when the ocean-land interface is a vertical cliff ($\theta = 90°$); (2) when OB is horizontal ($\theta = 0°$). The latter case is the more realistic one since the bottom slope seldom exceeds 5 to 10°.

3. The Horizontal Interface. In this problem, equation (13) takes the form

$$\varphi(x,0) = 1, \qquad (x < 0). \tag{15}$$

It is convenient to introduce the Fourier transform of φ and, in fact, we shall find φ by invoking the Wiener-Hopf technique. We define

$$\bar{\varphi}(\xi,y) = \int_{-\infty}^{\infty} \varphi(x,y)e^{-i\xi x}\,dx, \tag{16}$$

and the transform of equation (12) becomes

$$\bar{\varphi}_{yy} - (\xi^2 + i\epsilon)\bar{\varphi} = 0. \tag{17}$$

Hence

$$\bar{\varphi} = A(\xi)e^{y\sqrt{\xi^2+i\epsilon}}. \tag{18}$$

Now we can define $h_1(x)$ and $h_2(x)$ by

$$h_1(x) = \lim_{a\to 0}\begin{cases} e^{ax}, & (x < 0), \\ 0, & (x > 0), \end{cases}$$

$$h_2(x) = \begin{cases} 0, & (x < 0), \\ ?, & (x > 0), \end{cases}$$

and $h_1(x) + h_2(x) = \varphi(x,0)$. However,

$$A(\xi) = \bar{\varphi}(\xi,0) = \overline{h_1}(\xi) + \overline{h_2}(\xi).$$

It can be ascertained that

$$\overline{h_1}(\xi) = \frac{1}{a - i\xi} \tag{19}$$

and that $\overline{h_2}(\xi)$, being the transform of a function which vanishes for $x < 0$, is regular in ξ whenever Im $(\xi) < \delta$, where δ is a nonnegative real number. We can now define

$$f(x) = \varphi_y(x,0) + i\varphi(x,0)$$

and

$$\begin{aligned} \bar{f}(\xi) &= (\sqrt{\xi^2 + i\epsilon} + i)A(\xi) \\ &= (\sqrt{\xi^2 + i\epsilon} + i)[\overline{h_1}(\xi) + \overline{h_2}(\xi)]. \end{aligned} \tag{20}$$

Because of our boundary conditions, $\bar{f}$ (like $h_1(\xi)$) is analytic in the upper half ξ-plane. We can anticipate, in fact, that each function of ξ appearing in equation (20) will be regular in a horizontal strip lying below $i(\epsilon/2)^{\frac{1}{2}}$ and above the smaller of $-i(\epsilon/2)^{\frac{1}{2}}$ and $-ia$. In this strip, then, equation (20) is a meaningful equation. We now anticipate that $G(\xi) = \sqrt{\xi^2 + i\epsilon} + i$ can be written

$$G(\xi) = G_+(\xi)G_-(\xi),$$

where $G_+(\xi)$ is analytic and has no zeros above $\xi = -i(\epsilon/2)^{\frac{1}{2}}$ and where $G_-(\xi)$ has the same properties below $\xi = i(\epsilon/2)^{\frac{1}{2}}$. We can write equation (20) in the form

$$(21) \qquad \frac{\bar{f}(\xi)}{G_+(\xi)} - G_-(-ia)\overline{h_1}(\xi) = \overline{h_2}(\xi)G_-(\xi) + [G_-(\xi) - G_-(-ia)]\overline{h_1}(\xi).$$

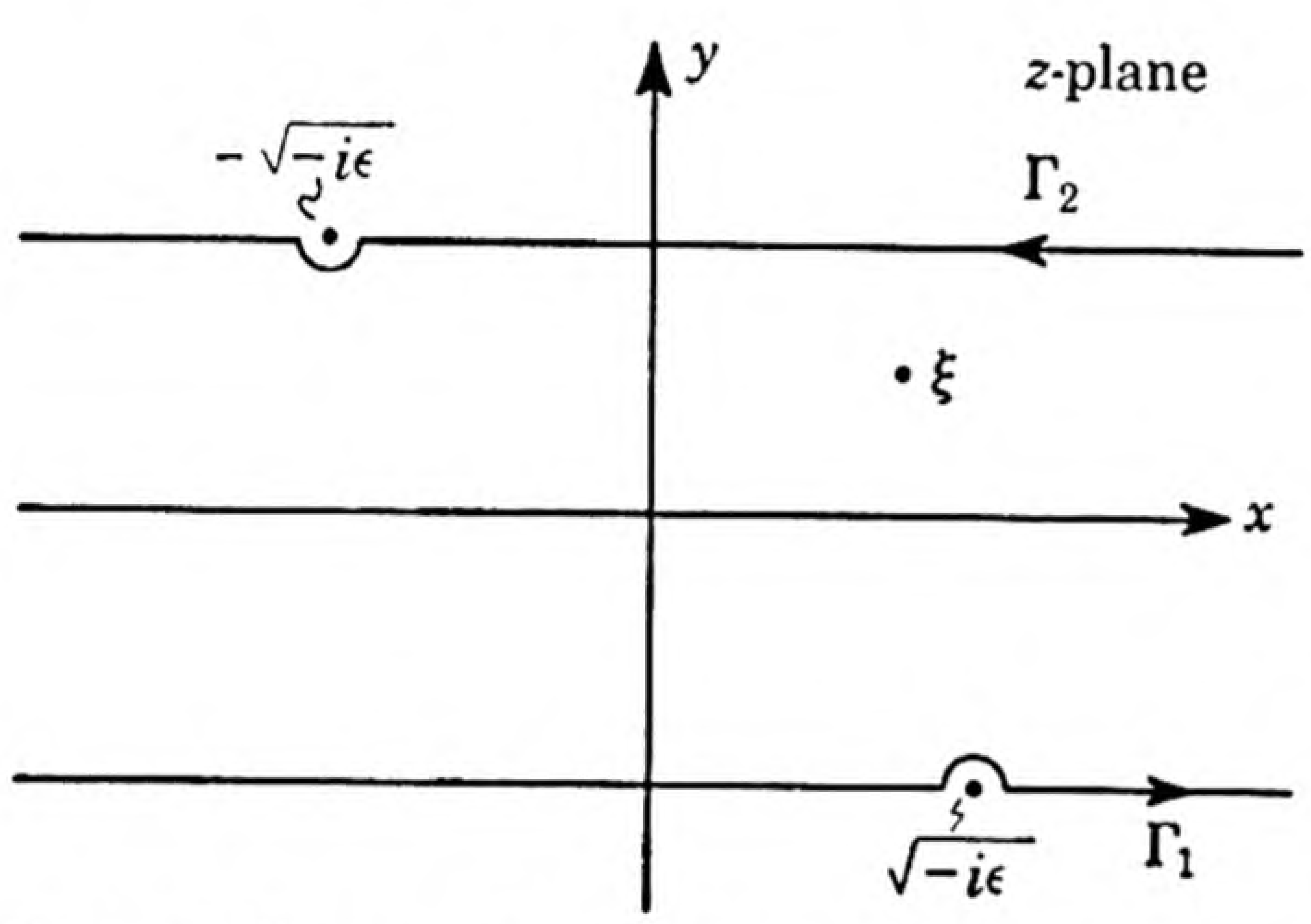


FIG. 2. Integration contour for equation (23).

The left side is analytic in the upper halt plane, the right side in the lower, and, by the usual analytic-continuation arguments, we may associate each side with an entire function. Since, as $\xi \to \infty$, both $\overline{h_1}(\xi) \to 0$ and $1/G_+(\xi) \to 0$, and since $\bar{f}(\xi)$ should be bounded at ∞, this entire function must be identically zero. Thus,

$$\overline{h_2}(\xi) = \left[\frac{G_-(-ia)}{G_-(\xi)} - 1\right]\overline{h_1}(\xi).$$

The free-surface level[2] is then given by

$$(22) \qquad h_2(x) = \frac{i}{2\pi}\int_{-\infty}^{\infty} \xi^{-1}\left[\frac{G_-(0)}{G_-(\xi)} - 1\right] e^{i\xi x}\, d\xi,$$

when we go to the limit $a = 0$. We cannot evaluate this, of course, until we find $G_-(\xi)$.

4. The Factorization of $G(\xi)$. We shall use a well-known technique [2] to obtain $G_-(\xi)/G_-(0)$. We note that

$$(23) \qquad \ln G(\xi) = \ln G_+(\xi) + \ln G_-(\xi) = \frac{1}{2\pi i}\int_\Gamma \frac{\ln G(z)}{z - \xi}\, dz$$

for all points in Γ, provided $\Gamma = \Gamma_1 + \Gamma_2$ is as indicated in Fig. 2. We may now define G_+ to be that contribution of equation (23) associated with Γ_1 and G_- to be that associated with Γ_2. It is clear that the appropriate analytic continuations of these functions into their respective half planes of definition are automatic. We want to compute, then,

$$H(\xi) = \ln G_-(\xi) = -\frac{1}{2\pi i}\int_{-\infty+i\sqrt{\epsilon/2}}^{+\infty+i\sqrt{\epsilon/2}} \ln [G(z)](z - \xi)^{-1}\, dz,$$

[2] Corresponding to unit tidal amplitude.

and, for Im $\xi < 0$, we may write

$$H(\xi) = -\frac{1}{2\pi i}\int_{-\infty}^{\infty} \ln\,[(z^2 + i\epsilon)^{\frac{1}{2}} + i](z - \xi)^{-1}\,dz. \tag{24}$$

If we differentiate once with regard to ξ and integrate by parts, we obtain a readily integrable formula for $H'(\xi)$. In fact, the result is

$$H'(z) = \frac{1}{2(z - i\beta)} + \frac{1}{2(z^2 + \beta^2)} \left[\beta \ln \frac{\beta + 1}{\beta - 1} - \frac{z}{(z^2 + \alpha^2)^{\frac{1}{2}}} \ln \frac{z + (z^2 + \alpha^2)^{\frac{1}{2}}}{z - (z^2 + \alpha^2)^{\frac{1}{2}}}\right], \tag{25}$$

where the branches must be chosen so that G'_- is regular in the lower half plane. In this formula $\alpha^2 = i\epsilon$, $\beta^2 = 1 + i\epsilon$. In the limit $\epsilon \to 0$, we have

$$H'(\xi) \to +\frac{1}{2}(\xi - i)^{-1} - \frac{1}{\pi}\frac{\ln i\xi}{\xi^2 + 1}. \tag{26}$$

The quantity $G_-(0)/G_-(\xi)$ is defined by

$$\frac{G_-(0)}{G_-(\xi)} = -\int_0^{\xi} H'(z)\,dz, \tag{27}$$

where H' is given by equation (25) for the compressible situation and by equation (26) for the incompressible case. Equation (22) now yields

$$h_2(x) = \frac{i}{2\pi}\int_{-\infty}^{\infty} \xi^{-1}\left[\exp\left(-\int_0^{\xi} H'(z)\,dz + i\xi x\right) - 1\right] d\xi. \tag{28}$$

The evaluation of this integral will be discussed in Sec. 6.

5. The Vertical-cliff Problem. If we now consider the situation where the land-ocean interface is vertical, one must find the solution of

$$\Delta\varphi = i\epsilon\varphi, \qquad (\text{in } x > 0, y < 0). \tag{29}$$

$$\varphi(0,y) = 1, \qquad (y < 0). \tag{30}$$

$$\varphi_y(x,0) + i\varphi(x,0) = 0, \qquad (x > 0). \tag{31}$$

It is convenient to define $\psi(x,y) = \varphi_y + i\varphi$. The function ψ must always obey equation (29), and the boundary conditions on ψ become

$$\psi(0,y) = i, \qquad (y < 0), \tag{32}$$

and

$$\psi(x,0) = 0, \qquad (y > 0). \tag{33}$$

We can now extend the domain on which ψ is defined to the entire right half plane and replace the present problem by that of finding $\psi(x,y)$, which obeys (15) and the condition

$$\psi(0,y) = \begin{cases} i = \lim\limits_{a\to 0} ie^{ay}, & (y < 0), \\ -i = \lim\limits_{a\to 0} -ie^{-ay}, & (y > 0). \end{cases} \tag{34}$$

Since ψ will be odd in y, equation (33) will be automatically satisfied. The solution is now most readily obtained by using Fourier transforms again. We define $\bar{\psi}(x,\eta) = \int_{-\infty}^{\infty} \psi(x,y)e^{-i\eta y}\, dy$. Multiplying equation (29) by $e^{-i\eta y}$ and integrating yields

$$\bar{\psi}_{xx} - (\eta^2 + i\epsilon)\bar{\psi} = 0$$

so that

$$\bar{\psi} = A(\eta) \exp\,(-\sqrt{\eta^2 + i\epsilon}\, x).$$

The function A must be obtained from the boundary conditions, *i.e.*,

$$A = \bar{\psi}(0,\eta) = -2\eta(\eta^2 + a^2)^{-1}.$$

Finally, since $\bar{\psi} = i(\eta + 1)\bar{\varphi}$, we have

$$(35) \qquad \varphi = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{\varphi}e^{i\eta y}\, d\eta = \frac{i}{2\pi}\int_{-\infty}^{\infty} \frac{2\eta \exp\,(i\eta y - \sqrt{\eta^2 + i\epsilon}\, x)}{(\eta^2 + a^2)(\eta + 1)}\, d\eta.$$

In particular, on $y = 0$, we have

$$(36) \qquad \phi(x,0) = \exp\,(-\beta x) - \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{ze^{ixz}}{\sqrt{z^2 + \alpha^2}(z^2 + \alpha^2 + 1)}\, dz$$
$$= \exp\,(-\beta x) + \frac{2i}{\pi}\left\{\left(\operatorname{arcsinh}\frac{1}{\alpha}\right)\sinh \beta x - \int_0^x \cosh\,[\beta(x - \tau)]K_0(\alpha\tau)\, d\tau\right\},$$

where K_0 is the modified Hankel function and $\beta = (1 + i\epsilon)^{\frac{1}{2}}$. Finally, in the limit $\epsilon \to 0$, we note that

$$\operatorname{arcsinh}\frac{1}{\alpha} \sim \ln\,(2\epsilon^{-\frac{1}{2}}) - \frac{i\pi}{4}$$

and

$$K_0(\alpha,\tau) \sim -\ln\frac{\alpha}{2} - \gamma - \ln \tau.$$

Thus, for $\epsilon \to 0$,

$$\varphi(x,0) = e^{-x} + \frac{2i}{\pi}\int_0^x (\gamma + \ln \tau)\cosh\,(x - \tau)\, d\tau,$$

where γ is Euler's constant. This function, $\varphi(x,0)$, behaves asymptotically like $-2i/\pi x$.

6. The Results. We shall reduce the integral implied by equations (26) and (28) and those given by equation (36) to a form convenient for computation. For the horizontal-interface problem with $\alpha = 0$ (the incompressible case)

$$(37) \qquad \varphi(x,0) = \frac{i}{2\pi}\int_{-\infty}^{\infty} \frac{\left[\exp\left\{\pi^{-1}\int_0^{\xi}[\ln iz/(z^2 + 1)]\, dz + i\xi x\right\} - 1\right] dz}{\sqrt{1 + i\xi}}$$
$$= 2^{-\frac{1}{2}}e^{-x-i\pi/8} + \pi^{-1}e^{-5\pi i/8}\int_0^1 \frac{e^{-xy} - y^{-\frac{1}{2}}e^{-x/y}}{(1 + y)^{\frac{1}{2}}(1 - y)}\cos u\, dy$$
$$- \pi^{-1}e^{-i\pi/8}\int_0^1 \frac{e^{-xy} + y^{-\frac{1}{2}}e^{-x/y}}{(1 + y)^{\frac{1}{2}}(1 - y)}\sin u\, dy,$$

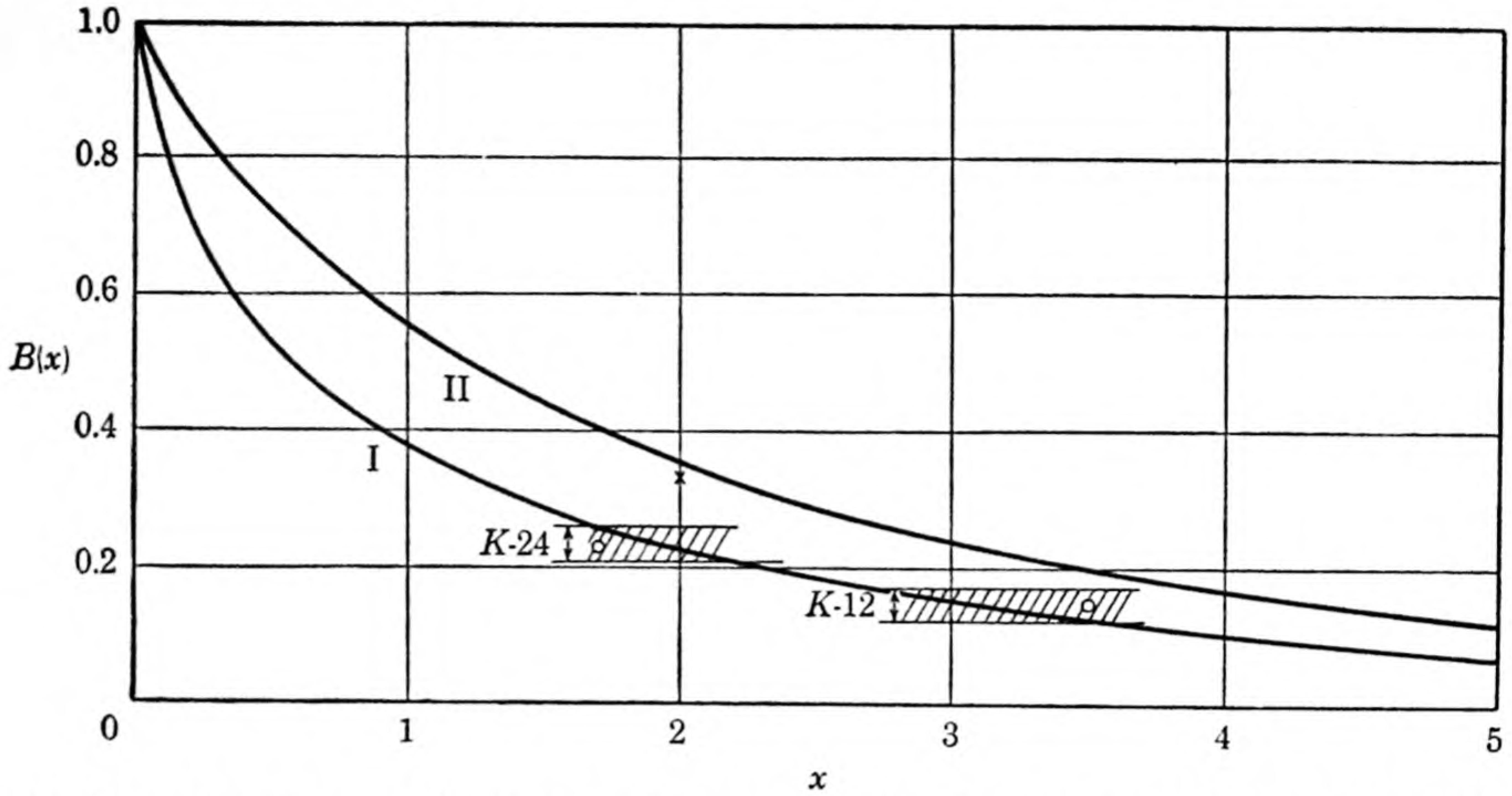


FIG. 3. Ground-water amplitude vs. distance inland. Curve I, incompressible horizontal interface. Curve II, incompressible vertical interface. x, $\epsilon = 0.01$ (vertical interface). O, average of observed amplitudes. K-12, 12-hr. tide observation range at Kuau. K-24, 24-hr. tide observation range at Kuau.

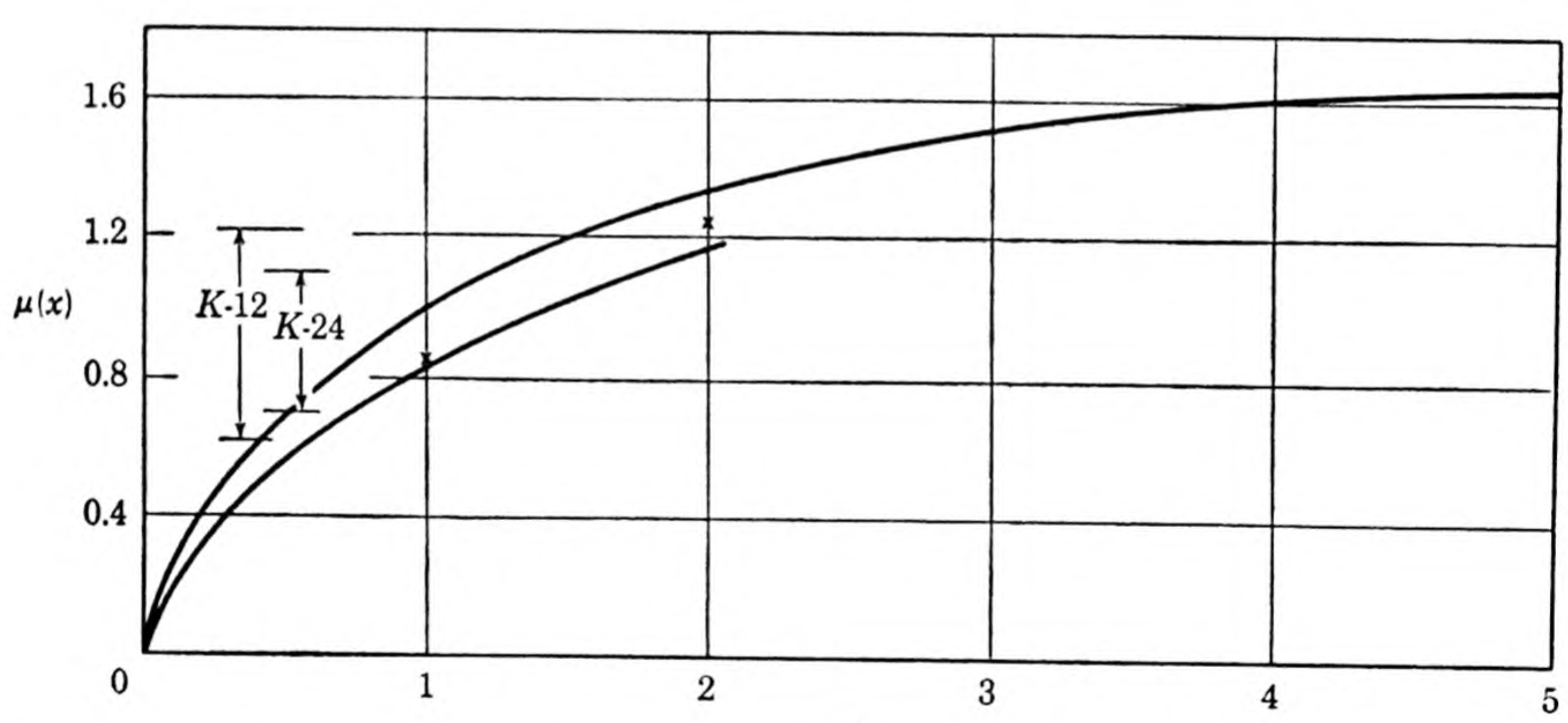


FIG. 4. Ground-water phase lag vs. distance inland. Curve I, incompressible horizontal interface. Curve II, incompressible vertical interface. x, $\epsilon = 0.01$ (vertical interface). K-12, 12-hr. tide observation range at well. K-24, 24-hr. tide observation range at well.

where $u = \pi^{-1} \int_v^1 [\ln v/(1 - v^2)]\, dv$. The numerical integration of this latter form yields the results depicted in Figs. 3 and 4.

The vertical-cliff result as given by equation (36) can also be written

$$\varphi(x,0) = e^{-x\sqrt{1+a^2}} - \frac{2i}{\pi} \int_0^1 \frac{e^{-x\sqrt{y^2+a^2}} - e^{-x/\sqrt{y^2+a^2}}}{1 - y^2}\, dy. \tag{38}$$

Asymptotically these functions [*i.e.*, the $\varphi(x,0)$ of equations (37) and (38)] behave like $-i/\pi x$ and $-2i/\pi x$, respectively, for $\alpha = 0$. The results of the

numerical integration of this equation are also plotted in Figs. 3 and 4, both for $\epsilon = 0$ and $\epsilon = 0.01$. The notation on these curves is given by

$$\varphi(x,0) = B(x)e^{-i\mu(x)}. \tag{39}$$

The experimental observations which are available are scanty, and the spread of the data is great. The observations at the Kuau well [1] are also indicated in Figs. 3 and 4. Note that the phase-lag observations are not easily interpreted but that the amplitude decay is reasonably consistent. Since, for a given well the "distance" x is proportional to the frequency, the 12- and 24-hr. tides should correspond to x values having the ratio 2. Choosing the location of these data points on this basis, we find that this location is consistent with a permeability value of 5×10^{-6} cm.^{2} and a value of ϵ of order 0.01. This permeability is consistent with other estimated values for this volcanic material. The observed phase lag, however, is much too low to be consistent with our theory. Furthermore, as is indicated by the vertical-cliff computations, the compressibility correction would lead to a greater discrepancy—not a smaller one. Thus, the phase-lag results are somewhat disappointing although the amplitude decay is consistent with our theory. This is not too surprising since the nonhomogeneity of the geologic setup can easily upset the validity of the results. Nevertheless, it is difficult to explain the very low phase-lag observations and especially the fact that this phase lag seems relatively independent of frequency.

Finally, it seems of some interest to record the asymptotic behavior of the function defined by equation (28) for the compressible case. One can show that

$$h_2(x) \approx -\frac{i}{2\pi}\left[K_1(\alpha x) - \int_{\alpha x}^{\infty} K_2(\tau)\,d\tau\right] - \frac{\alpha^2 \ln \alpha}{2\pi^2}\,[K_1'(\alpha x) + K_0(\alpha x)] + \cdots,$$

where K_ν is the Hankel function of imaginary argument. However, this series is an asymptotic representation in $1/\alpha$ rather than in x, and the function is well represented only when $x \gg |\ln \alpha|$. That is to say, it is useful only in a range of x where the compressibility is important. Since $|\alpha| = 0(0.1)$ and $x = 0(1)$ in the cases of interest to us, these results are not particularly pertinent to the present problem.

Bibliography

1. D. Cox and W. Munk, *The diffusion of tides into the volcanic rock of Hawaii: a method for estimating subsurface permeability* (in press).

2. H. Levine and J. Schwinger, *On the radiation of sound from an unflanged circular pipe*, Phys. Rev. vol. 73 (1948) p. 4.

Harvard University,
Cambridge, Mass.
Scripps Institution of Oceanography, University of California,
La Jolla, Calif.

SOME REMARKS ON RADIATION CONDITIONS

BY

J. J. STOKER

In linear wave-propagation problems for the so-called steady state (to be understood here as a motion that is simple harmonic in the time) in unbounded domains it is in general not possible to characterize uniquely the solutions having the desired physical characteristics by imposing only boundedness conditions at infinity. It is, in fact, necessary to impose sharper conditions. In the simplest case in which the medium is such as to include a full neighborhood of the point at infinity that is in addition made up of homogeneous matter, the correct radiation condition is not difficult to guess. It is simply that the wave at infinity behaves like an outgoing spherical wave from an oscillatory point source, and such a condition is what is commonly called the radiation, or Sommerfeld, condition. Among other things this condition precludes the possibility that there might be an incoming wave generated at infinity—which, if not ruled out, would make a unique solution of the steady-state problem impossible.

If the refracting or reflecting obstacles to the propagation of waves happen to extend to infinity—for example, if a rigid reflecting wall should happen to go to infinity—it is by no means clear a priori what conditions should be imposed at infinity in order to ensure the uniqueness of a steady-state solution having appropriate properties otherwise.[1] The point of view to be presented here is *that the difficulty arises because the steady-state problem is an unnatural problem in mechanics and that, in principle at least, one should rather formulate and solve an appropriate initial-value problem and then find the solution of the steady-state problem by making a passage to the limit in allowing the time to tend to infinity.*[2]

The steady-state problem is unnatural because one makes a hypothesis about the motion that holds for all time, while Newtonian mechanics is basically concerned with the *prediction*—in a unique way, furthermore—of the motion of a mechanical system from given initial conditions. Of course, in mechanics of continua that are unbounded it is necessary to impose conditions at infinity (in the space variables, that is) not derivable directly from Newton's laws, but for the initial-value problem it should suffice in general to impose only boundedness conditions at infinity.

[1] For a treatment of the radiation condition in such cases see F. Rellich, *Über das asymptotische Verhalten der Lösungen von $\Delta u + u = 0$ in unendlichen Gebieten*, Jber. Deutschen. Math. Verein vol. 53 (1943). Also see F. John, *On the motion of floating bodies* II, Comm. Pure Appl. Math. vol. 3 (1950) p. 54.

[2] The formulation of the usual radiation condition is doubtlessly motivated by an instinctive consideration of the same sort of hypothesis combined with the feeling that a homogeneous medium at infinity will have no power to reflect anything back to the finite region.

If one wished to be daring one might, on the basis of these remarks, formulate the following general method of obtaining the appropriate radiation condition for the steady-state problem: Consider any convenient problem in which the part of the domain outside a large sphere is maintained intact and initially at rest. (In other words, one might in particular feel free to modify in any convenient way any *bounded* part of the medium.) Next solve the initial-value problem for an oscillatory point source placed at any convenient point. A passage to the limit should then be made in allowing the time t to approach infinity, and after that the space variables should be allowed to approach infinity. The behavior at the far-distant portions of the domain should then furnish the appropriate radiation conditions independent of the constitution of the finite part of the domain. It might be worth pointing out specifically that this is a case in which the order of the two limit processes cannot be interchanged: obviously, if the time t were first held fixed while a space variable tends to infinity, the result would be that the motion would vanish at infinity, because of the finite propagation speed for disturbances, and no radiation conditions could be obtained.

The writer would not have set down these remarks—which are of a character so obvious that they must also have occurred to many others—if it were not for two considerations. Every reader will doubtlessly have said to himself, "That is all very well in principle, but will it not be prohibitively difficult to carry out the solution of the initial-value problem and to make the subsequent passages to the limit?" In general, such misgivings are probably all too well founded. However, the writer has come upon an interesting special case in which (1) the indicated program can be carried out in an elementary way in all detail and (2) it is slightly easier to obtain an integral representation for the solution of the initial-value problem than it is to solve the steady-state problem with the Sommerfeld condition imposed.

The problem in question concerns the propagation of surface-gravity waves in an infinite ocean. We consider only irrotational incompressible flow and assume further that the motion is a small oscillation in the neighborhood of the rest position of equilibrium; a linearized theory can then be obtained. We restrict ourselves to two-dimensional motion in an xy-plane, with the y-axis taken vertically upward and the x-axis in the originally undisturbed horizontal free surface. Under these assumptions there exists a velocity potential $\phi(x,y,t)$ which is a harmonic function in the lower half plane:

$$\phi_{xx} + \phi_{yy} = 0, \qquad (y < 0, t > 0). \tag{1}$$

(Here and in what follows subscripts denote partial derivatives.) The free-surface boundary conditions (cf. Lamb, *Hydrodynamics*, p. 364) are

$$\left.\begin{aligned} -\phi_y + \eta_t &= 0, && (2)\\ \phi_t + g\eta &= -\frac{1}{\rho}p, && (3)\end{aligned}\right\} \qquad (\text{for } y = 0, t > 0).$$

In these equations $\eta = \eta(x,t)$ represents the vertical displacement of the free surface measured from the x-axis, and $p = p(x,t)$ represents the pressure applied

on the free surface. We suppose that ϕ and its first and second derivatives tend to zero at ∞ for any given time t—in fact that they tend to zero in such a way that Fourier transforms exist—but we do not, in accordance with our discussion above, make any assumptions about the behavior of our functions as $t \to \infty$. At the time $t = 0$ we prescribe the following initial conditions,

$$\phi(x,0,0) = \phi_t(x,0,0) = p(x,0) = 0, \tag{4}$$

which imply that the free surface is initially at rest in its horizontal equilibrium position.

In what follows we consider only the special case in which the surface pressure $p(x,t)$ is given by

$$p(x,t) = \delta(x)e^{i\omega t}, \qquad (t > 0). \tag{5}$$

Here $\delta(x)$ is the Dirac δ-function; in other words, an oscillatory pressure point is applied at the origin on the free surface and maintained there for all time. By inserting this expression for p in (3) and eliminating the quantity η by making ue of (2) the free-surface condition is obtained in the forms

$$g\phi_y + \phi_{tt} = -\frac{i\omega}{\rho}\,\delta(x)e^{i\omega t}, \qquad (t > 0). \tag{6}$$

Our problem now consists in finding a solution $\phi(x,y,t)$ of (1) which behaves properly at ∞ and which satisfies the free-surface condition (6) and the initial conditions (4).[3]

We proceed to solve the initial-value problem by making use of the Fourier transform applied to the variable x, or rather, since the problem is evidently symmetrical with respect to the y-axis, we may use the cosine transform.[4] The result of transforming (1) is

$$-s^2\bar{\phi} + \bar{\phi}_{yy} = 0, \tag{7}$$

in which $\bar{\phi}(s,y,t)$ is the cosine transform of $\phi(x,y,t)$ and use has been made of the fact that $\phi_x = 0$ for $x = 0$, $y < 0$, and $\phi_x \to 0$ as $x \to \infty$. The bounded solutions of (7) for $y < 0$, $s > 0$ are all of the form

$$\bar{\phi}(s,y,t) = A(s,t)e^{sy}. \tag{8}$$

The cosine transform is now applied to the boundary condition (6), with the result

[3] The steady-state problem is formulated as follows: One sets $\phi = \psi(x,y)e^{i\omega t}$. The time factor $e^{i\omega t}$ can be canceled from (1) and (6), and the harmonic function then satisfies the free-surface condition $g\psi_y - \omega^2\psi = -\frac{i\omega}{\rho}\,\delta(x)$. In addition one would postulate that the motion should have the character of an outgoing progressing wave at $x = -\infty$ and $x = +\infty$. This problem was solved long ago by H. Lamb, *On deep water waves*, Proc. London Math. Soc. Ser. 2 vol. 2 (1904) pp. 371–400 (see pp. 388–389). It is a curious fact that this problem, which involves the determination of a *harmonic* function and not a solution of a wave equation, nevertheless has not a unique solution of the desired type if only boundedness conditions are imposed at ∞.

[4] The same general approach to problems of this type is used in the book by I. N. Sneddon, *Fourier transforms*, McGraw-Hill Book Company, Inc., New York, 1951, p. 278.

(9) $$g\bar{\phi}_y + \bar{\phi}_{tt} = -\frac{1}{\sqrt{2\pi}}\frac{i\omega}{\rho}e^{i\omega t}, \qquad (\text{for } y = 0),$$

and on substitution of $\bar{\phi}(s,0,t)$ from (8) we find

(10) $$A_{tt} + gsA = -\frac{1}{\sqrt{2\pi}}\frac{i\omega}{\rho}e^{i\omega t}.$$

The initial conditions (4) now furnish for $A(s,t)$ the conditions

(11) $$A(s,0) = A_t(s,0) = 0.$$

The solution of (10) subject to the initial conditions (11) is

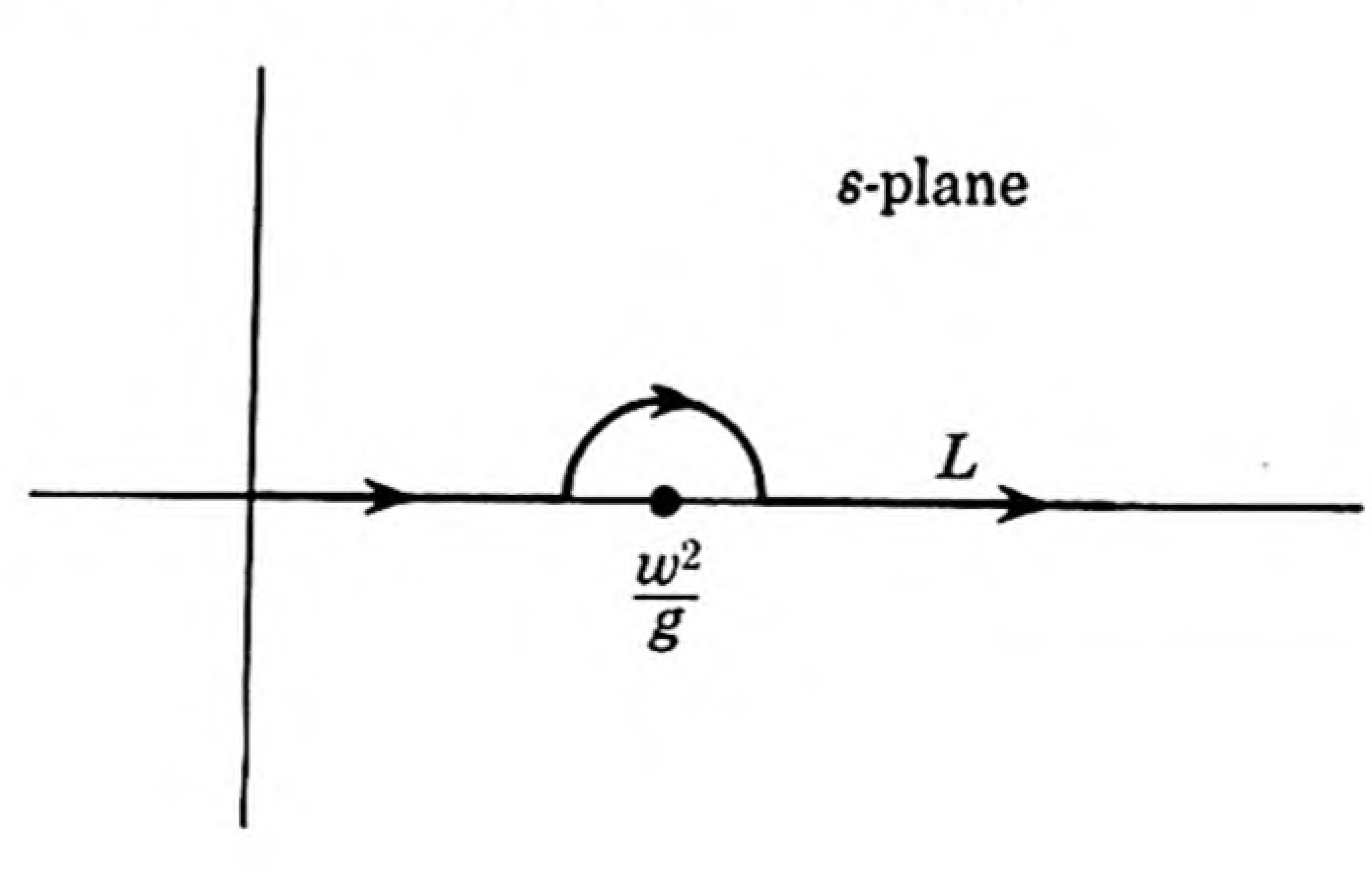
(12) $$A(s,t) = -\frac{1}{\sqrt{2\pi}}\frac{i\omega}{\rho}\int_0^t \frac{e^{i\omega(t-\tau)}}{\sqrt{gs}}\sin\sqrt{gs}\,\tau\,d\tau.$$

Finally, we insert the last expression for $A(s,t)$ in (8) and apply the inverse transform to obtain the following integral representation for our solution $\phi(x,y,t)$:

(13) $$\phi(x,y,t) = -\frac{i\omega}{\rho\pi}\int_0^\infty e^{sy}\cos sx \int_0^t \frac{e^{i\omega(t-\tau)}}{\sqrt{gs}}\sin\sqrt{gs}\,\tau\,d\tau\,ds.$$

This solution is also seen to be the unique solution of our problem. Our object is to study the behavior of this solution as $t \to \infty$.

Since y is negative (we do not discuss here the limit as $y \to 0$, that is, the behavior on the free surface), the integral with respect to s converges well and there is no singularity on the positive real axis of the complex s-plane. However, the passage to the limit $t \to \infty$ is more readily carried out by writing the solution in a different form in which a singularity—a pole, in fact—then appears on the real axis of the s-plane. (It seems, indeed, likely that such an occurrence would be the rule in any considerations of the present kind since the limit function as $t \to \infty$ would not usually be a function having a Fourier transform, and one could expect that the limit function would somehow appear as a contribution in the form of a residue at a pole.) It is convenient—and of course legitimate by Cauchy's theorem—to replace the path of integration along the s-axis in the s-plane by a path L indicated in the accompanying figure. The path L lies on the positive real axis except for a semicircle in the upper half plane centered at the point $s = \omega^2/g$.



We now replace $\sin \sqrt{gs}\,\tau$ in (13) by exponentials and carry out the integration on τ to obtain[5]

$$(14) \quad \phi(x,y,t) = -\frac{i\omega e^{i\omega t}}{\pi\rho}\int_L e^{sy}\cos sx \left[-\frac{1}{2\sqrt{gs}}\frac{e^{i(\sqrt{gs}-\omega)t}}{\sqrt{gs}-\omega} - \frac{1}{2\sqrt{gs}}\frac{e^{-i(\sqrt{gs}+\omega)t}}{\sqrt{gs}+\omega} + \frac{1}{gs-\omega^2}\right] ds.$$

We wish now to consider the three items in the bracket separately, and as we see, two of them do indeed have a singularity at $s = \omega^2/g$ which is bypassed through our choice of the path L. The first two items are rather obviously the result of the initial conditions and hence could be expected to provide transients which die out as $t \to \infty$. This is in fact the case, as can be seen easily in the following way: That branch of $\sqrt{s}$ is taken which is positive on the positive real axis, and we operate always in the right half plane. If, in addition, s is in the upper half plane, it follows that $i(\sqrt{gs} \pm \omega)$ has its real part negative (ω being real). Consider now the contribution furnished by the first item in the square brackets. Since the exponential has a negative real part on the semicircular portion of the path L, it is clear that as $t \to +\infty$, this part of the path makes a contribution that tends to zero as $t \to \infty$. The remaining portions of L, which lie on the real axis, are then readily seen to make contributions which die out like $1/t$: this can be seen easily by integration by parts, for example, or by application of known results about Fourier transforms. The second item in the square brackets has no singularity on the real axis, so that the path L can be taken entirely on the real axis; thus, in accordance with the remarks just made concerning the similar situation for the first item, it is clear that this contribution also dies out like $1/t$. Thus for large t we obtain the following asymptotic representation for ϕ:

$$(15) \qquad \phi(x,y,t) \simeq -\frac{i\omega}{\pi\rho}e^{i\omega t}\int_L \frac{e^{sy}\cos sx}{gs-\omega^2}\,ds.$$

Actually, the right-hand side is the solution of the steady-state problem—as obtained, for example, in the paper of Lamb cited previously—when the condition at ∞ is the radiation condition stating that ϕ behaves like an outgoing progressing wave. The steady-state solution furnished by (15) is a little more awkward to obtain directly through use of the radiation condition—at least by methods known to the author—than obtaining solution (13) of the initial-value problem. In particular, the asymptotic behavior of an integral representation must be investigated in this case also before the radiation condition can be used. Thus we have verified in this special case that the radiation condition can be replaced by boundedness conditions (in the space variables, that is) if one treats an appropriate initial-value problem instead of the steady-state problem.

[5] The argument which follows was suggested to the writer by his colleague A. S. Peters.

Even though not strictly necessary—since (15) is known to furnish the desired steady-state solution—it is perhaps of interest to show directly that the right-hand side of (15) has the behavior one expects for an outgoing progressing wave when $x \to +\infty$. The procedure is the same as that used above in discussing (14): The factor $\cos sx$ is replaced by exponentials to obtain

$$(16) \qquad \phi(x,y,t) \simeq \frac{\omega}{\rho} e^{i\omega t} \left(\frac{1}{2\pi i} \int_L \frac{e^{sy} e^{isx}}{gs - \omega^2} ds + \frac{1}{2\pi i} \int_L \frac{e^{sy} e^{-isx}}{gs - \omega^2} ds \right).$$

By the same argument as above one sees that the first integral makes a contribution that tends to zero as $x \to +\infty$. The second integral is treated by deforming the path L over the pole $s = \omega^2/g$ into a path M which consists of the positive real axis except for a semicircle *in the lower half plane*. The contribution of the second integral then consists of the residue at the pole plus the integral over the path M. But the contribution of the latter integral is, once more, seen to tend to zero as $x \to +\infty$ because of the factor e^{-isx}. Thus $\phi(x,y,t)$ behaves for large x as follows:

$$(17) \qquad \phi(x,y) \simeq -\frac{\omega}{g\rho} e^{(\omega^2/g)y} e^{-i[(\omega^2/g)x - \omega t]}.$$

This represents a progressing wave in the positive x-direction which, in addition, has the wavelength $2\pi g/\omega^2$ appropriate to a progressing sine wave with the frequency ω in water of infinite depth.

New York University,
New York, N.Y.

ON THE THEORY OF SCATTERING OF PLANE WAVES BY SOFT OBSTACLES[1]

BY

E. W. MONTROLL AND J. M. GREENBERG

1. Introduction. One of the most important physical processes under investigation today is the scattering of waves by various types of obstacles. In the realm of basic physics and physical chemistry, experimental scattering techniques are used to study the detailed character of atomic nuclei (and nuclear forces), the arrangement of atoms in simple molecules, and the sizes and shapes of large molecules such as the virus proteins. Observations of the scattering of light from distant stars are used to deduce the nature of interstellar dust.

The general procedure is to choose a model for the scatterer and to compare observed scattering patterns with those calculated on the basis of the model. If an observed pattern is in agreement with that deduced from the model, one has considerable evidence that it is a reasonable one. This approach is of course quite old, having been used by Lord Rayleigh in his theory of the color of the sky and by Rutherford in his theory of atomic structure.

Modern technology is also concerned with scattering processes. For example, the scattering of sound and radio waves is important in the fields of long-range communication and the application of radar and sonar.

Although scattering processes have been known and investigated for many years, the theory of the processes has usually lagged somewhat behind the demands made upon it. Even though a tremendous amount of work has been done recently to rectify this situation, there is still much to be done by applied mathematicians and physicists in this field.

The aim of this paper is to report on some progress that has been made by van de Hulst, Hart, Glauber, and the authors on the problem of scattering of plane waves by "soft obstacles." An obstacle is considered to be soft if the wavelength of the wave inside the scatterer does not differ much from that of the incident wave in the absence of the scatterer. The well-known approximation of Rayleigh, Gans, and Born in which the internal field of an obstacle is taken to be that which would have existed in the absence of the scatter is ordinarily used in the analysis of scattering by soft obstacles. Unfortunately, when the range of the scatterer is more than a few wavelengths of the incident radiation, this approximation becomes very poor.

2. Fundamental Equations. We shall introduce the fundamental equations needed for the development of scattering theory in this section. Although we restrict ourselves to scattering of scalar waves with "quantum-mechanical" boundary conditions—wave functions and their first derivatives continuous—

[1] This work was supported by the Office of Naval Research.

one can (not always without some effort) generalize the methods and results to vector waves and other boundary conditions. These boundary conditions correspond to the acoustical case in which the density of the scatterer is the same as that of the medium. They are the scalar analogue of the scattering of light by a dielectric.

The fundamental wave equation is

$$\nabla^2\psi + k^2(r)\psi = 0. \tag{1}$$

The quantity $|\psi|^2$ has several physical interpretations depending on the problem at hand. It is the particle density in the quantum theory, proportional to sound intensity in acoustics, and to light intensity in the corresponding vector equation of light scattering.

We let

$$\psi_0 = \exp ik_0(\mathbf{r} \cdot \mathbf{s}_0) \tag{2}$$

represent the wave function of an incident plane wave propagated in the direction of the unit vector $\mathbf{s}_0$. Its wavelength λ_0 is related to k_0 through

$$k_0 = \frac{2\pi}{\lambda_0} \tag{3}$$

In the quantum-mechanical case we have

$$k_0^2 = \frac{2mE}{\hbar^2}, \tag{3a}$$

$$k^2(r) = \frac{2m[E - V(r)]}{\hbar^2}, \tag{3b}$$

where E is the energy associated with an incident particle, m its mass, $V(r)$ the potential of the scattering field, and $\hbar$ = Planck's constant/2π. We shall frequently refer to

$$m(r) = \frac{k(r)}{k_0} \tag{3c}$$

as the relative index of refraction of scatterer to medium at the point r and to $m = k_1/k_0$ as the relative index of a uniform homogeneous scatterer to its medium.

If $\psi_s(r)$ is the wave function of the scattered field, the solution of (1) is the form

$$\psi(r) = \psi_s(r) + \exp ik_0(\mathbf{r} \cdot \mathbf{s}_0) \tag{4}$$

where $\psi_s(r) \to 0$ as $1/r$. Hence $\psi_s(r)$ satisfies the differential equation

$$\nabla^2\psi_s + k_0^2\psi_s = [k_0^2 - k^2(r)]\psi \tag{5}$$

and therefore the equivalent integral equation [1]

$$\psi_s(R) = \frac{1}{4\pi}\int \frac{\exp ik_0|\mathbf{R} - \mathbf{r}|}{|\mathbf{R} - \mathbf{r}|}[k^2(r) - k_0^2]\psi(r)\, d\tau; \tag{6}$$

the integration extends over all space.

A physical interpretation of (6) based on considering every infinitesimal element of the scatterer as a Rayleigh scatterer is suggestive in seeking an approximate solution of (6). The quantity $(\exp ik_0|\mathbf{R} - \mathbf{r}|)/4\pi|\mathbf{R} - \mathbf{r}|$ is proportional to the wave function of the radiation emitted from an infinitesimal sphere at the point $\mathbf{r}$. The term $k^2(r) - k_0^2$ expresses the difference between the scatterer and the medium at the point $\mathbf{r}$, and $\psi(r)$ is the internal field at the point $\mathbf{r}$ which excites the infinitesimal scatterer of volume $d\tau$. Hence (6) corresponds to a superposition of the scattered radiation of a large number of infinitesimal Rayleigh scatterers, each excited by the proper internal field at its point of location. Hence if on some physical basis we can make a good approximation to the internal wave function $\psi(r)$, we can substitute it into (6), integrate, and obtain an approximate scattered wave function.

The theory of scattering most often used, the Rayleigh [2]–Gans [3]–Born [4] approximation, is obtained by approximating $\psi(r)$ by the wave function associated with the incident wave in the absence of a scatterer. If in the quantum-mechanical case $|E| \gg |V(r)|$ so that $|k^2(r) - k_0^2|$ is small, this approximation is good unless the scattering center has a range larger than several wavelengths of the incident wave. When the range is large, the incident wave becomes more and more distorted with increasing distance of passage through the scattering field. An improved estimation of the internal field will be discussed in the remainder of this article.

When observations are made at great distances from the scatterer, (6) can be simplified to

$$\psi_s(R) \sim \frac{e^{ik_0R}}{4\pi R} \int \psi(r)[k^2(r) - k_0^2] \exp(-ik_0\mathbf{r} \cdot \mathbf{s}_1)\, d\tau \tag{7}$$

where $\mathbf{s}_1$ is a unit vector in the direction of the scattered wave. Although improved scattered wave functions can be obtained in principle by substituting an approximate internal wave function into (6) and iterating several times, it becomes an almost hopeless procedure in practice because $\psi_s(R)$ can seldom be computed for *all* $R > 0$ by (6). The asymptotic $\psi_s(R)$ as $R \to \infty$ is usually quite easy to derive from (7), but unfortunately even a second iteration for $\psi_s(R)$ requires $\psi(R)$ for all R.

Since the ingenious variational principal of Schwinger and Levine is essentially a second iteration process, it is difficult to apply in most cases.

3. Scattering by Soft Spheres. We consider a sphere to be soft when the internal propagation constant k_1 ($k_1 = k(r)$ for $r < a$, the sphere radius) is almost equal to the external constant k_0. In quantum-mechanical language a spherical obstacle corresponds to a square-well (or barrier) scattering potential. The soft sphere represents a well or barrier whose depth or height is small compared with the energy of the incident wave. The R.G.B. approximation applies very well to the scattering by a soft sphere whose radius is not too much larger than the wavelength of the incident wave. When several wavelengths

of the incident wave fit into the sphere the incident wave function is a bad approximation to the sphere's internal wave function.

The qualitative nature of the internal wave function of the sphere can be guessed by considering the analogous one-dimensional problem of the propagation of a plane wave through an infinite flat plate. Let the plane wave be propagated from left to right in Fig. 1. The wave will be partially reflected by the surface of the plate. The transmitted wave will also be partially reflected by and partially transmitted through the second surface of the plate.

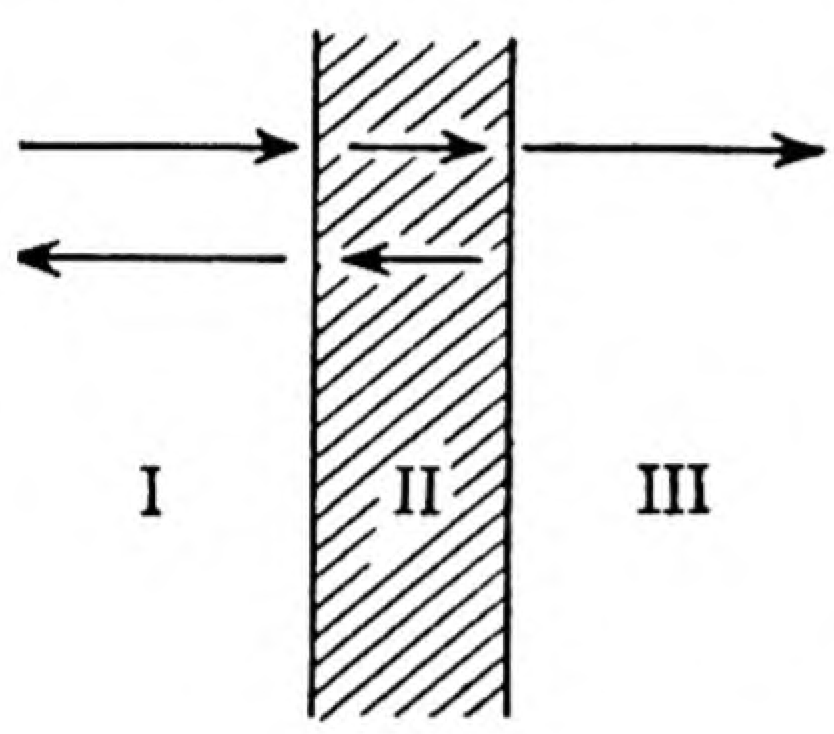


FIG. 1

Hence, the forms of the wave function for the three regions of Fig. 1 are:

$$\begin{aligned} \text{I. } \psi(x) &= e^{ixk_0} + Ae^{-ixk_0}, \\ \text{II. } \psi(x) &= Be^{ixk_1} + Ce^{-ixk_1}, \\ \text{III. } \psi(x) &= De^{ixk_0}, \end{aligned}$$

where the constants A, B, C, and D are to be determined from the boundary conditions $\psi(x)$ and $\psi'(x)$ continuous, at the boundaries of the plate.

One might expect the internal wave function of a soft sphere to be of the form of a linear combination of a transmitted and reflected wave with the propagation constant k_1 associated with the scatterer. This is indeed the case, as we shall show below.

The exact expressions for the internal and scattered wave functions of a spherical obstacle are easily derived from the fundamental wave equation

$$\nabla^2\psi + k_1^2\psi = 0, \qquad \left(\text{if } r < a,\ k_1 = \frac{2\pi}{\lambda_1}\right), \tag{8a}$$

$$\nabla^2\psi + k_0^2\psi = 0, \qquad \left(\text{if } r > a,\ k_0 = \frac{2\pi}{\lambda_0}\right) \tag{8b}$$

(λ_0 = wavelength of incident wave and λ_1 that of the corresponding wave in a medium composed of the material of the sphere) with the boundary conditions

$$\left.\begin{matrix} r\psi, \\ \dfrac{\partial(r\psi)}{\partial r}, \end{matrix}\right\} \qquad (\text{continuous at } r = a).$$

Let the incident wave be propagated in the z-direction. Then its wave function is

$$\psi_i = e^{izk_0} = e^{ik_0 r \cos\theta}. \tag{9}$$

It is easily shown in the usual manner [5–7] that the transmitted and scattered waves are given by

$$\psi_t = \sum_{n=0}^{\infty} A_{t,n} i^n (2n+1) j_n(rk_1) P_n(\cos\theta), \tag{10a}$$

$$\psi_s = \sum_{n=0}^{\infty} A_{s,n} i^n (2n+1) h_n(rk_0) P_n(\cos\theta), \tag{10b}$$

where

$$A_{t,n} = \frac{i/k_0a^2}{k_1h_n(ak_0)j_{n+1}(ak_1) - k_0j_n(ak_1)h_{n+1}(ak_0)}, \tag{11a}$$

$$A_{s,n} = -\frac{k_1j_n(ak_0)j_{n+1}(ak_1) - k_0j_n(ak_1)j_{n+1}(ak_0)}{k_1h_n(ak_0)j_{n+1}(ak_1) - k_0j_n(ak_1)h_{n+1}(ak_0)}. \tag{11b}$$

Then

$$\psi = \psi_i + \psi_s, \qquad \text{(if } r > a\text{)},$$
$$\psi = \psi_t, \qquad \text{(if } r < a\text{)}.$$

Here $z_n(u)$ is the spherical Bessel function

$$z_n(u) = \left(\frac{\pi}{2u}\right)^{\frac{1}{2}} Z_{n+\frac{1}{2}}(u),$$

where z and Z represents either j and J (Bessel) or h and H (Hankel) function.

The rapidity of convergence of (10a) and (10b) depends on the value of the ratio $x = 2\pi a/\lambda$. Debye [8] has shown that the number of significant terms in these series is approximately the first integer larger than x. This result can be observed empirically in scattering tables [9, 10]. The following argument gives a physical basis to Debye's theorem: If the series (10a) and (10b) are interpreted as the internal and scattered wave functions of a particle scattered by a force center of range a, each term in the series represents an angular momentum state of the particle. The lth state has an angular momentum $\hbar[l(l+1)]^{\frac{1}{2}}$. Classically those particles which are outside the range a of the scattering force are unaffected by it. The angular momentum with respect to the force center of a particle of linear momentum p at the boundary of the range of the sphere of interaction is pa. Hence, the largest l which can contribute significantly to the series (10) is defined by $pa = \hbar[l(l+1)]^{\frac{1}{2}}$. But a particle of momentum p corresponds to a wave of wavelength $\lambda = 2\pi\hbar/p$. Hence, l is related to λ and a by $2\pi a/\lambda = [l(l+1)]^{\frac{1}{2}}$, which for large l is Debye's result.

The Debye theorem suggests an approach to the approximation of the sums ψ_s and ψ_t. Let us suppose we can accurately approximate (11a) and (11b) by simpler expressions when $n < x$ so that the series for ψ_s and ψ_t can be summed with the approximate coefficients. If the approximate coefficients for $n > x$ are small, even though they might be in considerable error, Debye's theorem would indicate that the new summable series should be good representations of ψ_s and ψ_t.

We shall now attempt to simplify the coefficients (11a) and (11b) by applying the asymptotic formulas [11]

$$h_n(x) \cong (-i)^{n+1}x^{-1}e^{ix}, \tag{12a}$$

$$j_n(x) \cong x^{-1} \cos \frac{x - \pi(n+1)}{2}, \tag{12b}$$

which are valid when $x \gg n$. These formulas yield the following approximation to the denominators of (11a) and (11b):

$$(13) \qquad \frac{i(k_1 + k_0)}{2a^2 k_1 k_0} e^{-ia(k_1 - k_0)}[1 + (-)^n k e^{2iak_1}]; \qquad \left(k = \frac{k_1 - k_0}{k_1 + k_0}\right).$$

An especially interesting feature of this formula is that it reduces to the exact formula for the denominator as $k_1 \to k_0$ *for all integral* n's. Hence, a result based on (13) should be valid for soft spheres, and indeed it should be exact in the limit as $k_1 \to k_0$. Unfortunately, (13) is a very poor approximation to the denominators of (11a) and (11b) in the case of $|k|$ large. Then it yields coefficients to (10) which do not satisfy the Debye condition of being small when $n > 2\pi a/\lambda$.

When the approximate coefficients, obtained by employing (13) in (10a) and (10b), are used, both series can be summed by using well-known Bessel-function formulas. We first sum (10a) to obtain the internal wave function. For $|r| < a$ we have

$$(14a) \quad \psi_t(r) \simeq 2k_1 \frac{e^{ia(k_1 - k_0)}}{(k_1 + k_0)(1 - k^2 e^{4iak_1})} \sum_{n=0}^{\infty} [1 - (-)^n k e^{2iak_1}] i^n (2n + 1) j_n(rk_1) P_n(\cos\theta).$$

If we neglect the term proportional to k, the sum is exactly the spherical harmonic expansion of a plane wave propagated in the z-direction,

$$e^{ik_1 z} = \exp(ik_1 r \cos\theta).$$

The term proportional to k in the sum is the corresponding expansion for a plane wave propagated in the $-z$-direction. Hence

$$(14b) \quad \psi_t(r) \simeq \frac{2k_1 e^{ia(k_1 - k_0)}}{(k_1 + k_0)(1 - k^2 e^{4iak_1})} \{\exp[i(\mathbf{r} \cdot \mathbf{s}_0)k_1] - k e^{2iak_1} \exp[-i(\mathbf{r} \cdot \mathbf{s}_0)k_1]\}$$

so that the internal wave function corresponds to the superposition of two plane waves, one propagated in the direction of the incident wave and the other in the opposite direction. The vector $\mathbf{s}_0$ is a unit vector in the direction of the incident wave.

The scattered wave function can be computed with the appropriate denominator in $A_{s,n}$ by using the addition theorem for Bessel functions. However, as a preview of Sec. 5 we find $\psi_s(r)$ by substituting (14b) into the integral equation (7) with

$$k^2(r) - k_0^2 = \begin{cases} k_1^2 - k_0^2, & (r < a), \\ 0\ , & (r > a), \end{cases}$$

and, proceeding in the same manner as is done in the Born approximation, we obtain for large R

(15) $$\psi_s(R) \sim \frac{2k_1(k_1 - k_0)e^{iRk_0+ia(k_1-k_0)}}{2\pi R(1 - k^2e^{4iak_1})} (I_1 - ke^{2iak_1}I_2),$$

where

(15a) $$I_1 = \int_v \exp [i\mathbf{r} \cdot (k_1\mathbf{s}_0 - k_0\mathbf{s}_1)] \, d\tau$$

and

(15b) $$I_2 = \int_v \exp [-i\mathbf{r} \cdot (k_1\mathbf{s}_0 + k_0\mathbf{s}_1)] \, d\tau.$$

The integration extends over the volume of the obstacle.

To integrate I_1 we define γ as the angle between the fixed vector $(k_1\mathbf{s}_0 - k_0\mathbf{s}_1)$ and the variable vector r which ranges through the entire sphere. Let

(16a) $$\omega_1^2 = |k_1\mathbf{s}_0 - k_0\mathbf{s}_1|^2 = k_1^2 + k_0^2 - 2k_1k_0 \cos \theta,$$

where θ is the angle between s_0 and s_1. Then

(16b) $$\begin{aligned} I_1 &= 2\pi \int_0^a \int_0^\pi \exp (ir\omega_1 \cos \gamma)r^2 \sin \gamma \, d\gamma \, dr \\ &= 4\pi\omega_1^{-3} \int_0^{a\omega_1} u \sin u \, du \\ &= 4\pi a^3 \left(\frac{\pi}{2}\right)^{\frac{1}{2}} (a\omega_1)^{-\frac{3}{2}}J_{\frac{3}{2}}(a\omega_1), \end{aligned}$$

where $J_{\frac{3}{2}}(x)$ is the Bessel function of order $\frac{3}{2}$. Similarly we have

(16c) $$I_2 = 4\pi a^3 \left(\frac{\pi}{2}\right)^{\frac{1}{2}} (av_1)^{-\frac{3}{2}}J_{\frac{3}{2}}(av_1),$$

where

(16d) $$v_1^2 = k_1^2 + k_0^2 - 2k_1k_0 \cos \theta.$$

By substituting (16b) and (16c) into (15) we can easily derive an approximate expression for the differential-scattering cross section $\sigma_s(\theta)$ which has the property that $\sigma_s(\theta) \, d\Omega$ is the fraction of incident energy associated with the incident plane wave which is scattered through a solid angle $d\Omega$ at θ. We find [7]

(17) $$\begin{aligned} \sigma_s(\theta) &= R^2|\psi_s|^2 \\ &= |A_s|^2 \left[\frac{J_{\frac{3}{2}}{}^2(a\omega_1)}{(a\omega_1)^3} - 2k \cos (2ak_1) \frac{J_{\frac{3}{2}}(a\omega_1)J_{\frac{3}{2}}(av_1)}{(a^2v_1\omega_1)_{\frac{3}{2}}} + k^2 \frac{J_{\frac{3}{2}}{}^2(av_1)}{(av_1)^3}\right] \end{aligned}$$

with

$$|A_s|^2 = \frac{4\pi k_1^2(k_1 - k_0)^2a^6}{1 + k^4 - 2k^2 \cos 4ak_1}.$$

This result is exactly that obtained by summing (10b) with the approximate denominator (13) used in (11b).

A similar result can be derived for the scattering cross section of an infinite circular cylinder. In the case of normal incidence of the plane wave to the cylinder one has [12]

$$\sigma_s(\theta) \sim \frac{2\pi a^4 k_1 (m-1)^2 k_0^2}{1+k^4+2k^2 \cos 4ak_1} \left\{ \frac{J_1^2(a\omega_1)}{(a\omega_1)^2} + 2k \sin 2ak_1 \frac{J_1(a\omega_1)J_1(av_1)}{a^2\omega_1 v_1} + \frac{k^2 J_1^2(av_1)}{(av_1)^2} \right\},$$

where ω_1 and v_1 are defined as before with θ as the polar angle about the cylinder in cylindrical coordinates.

A somewhat different analysis of the problem of scattering by soft spheres has been made by van de Hulst [13] (see especially pages 45 to 46). First he applies Huygens' principle to the problem. In the limit as $m = k_1/k_0 \to 1$ a negligible part of the incident radiation is reflected, and the refraction of the transmitted part causes only slight deviation in the direction of an incident ray. Hence, in quantum-mechanical language, a particle that passes through the sphere changes neither its direction nor intensity. It, however, changes its phase by an amount proportional to the length of the chord along which it passes through the soft sphere. As will be discussed at greater length in Sec. 6, the change in phase along a chord whose closest distance to the center of the sphere is u is

$$\begin{aligned} \chi(u) &= 2k_0(m-1)(a^2-u^2)^{\frac{1}{2}} \qquad (u \le a) \\ &= 0, \qquad\qquad\qquad\qquad\quad (u \ge a). \end{aligned}$$

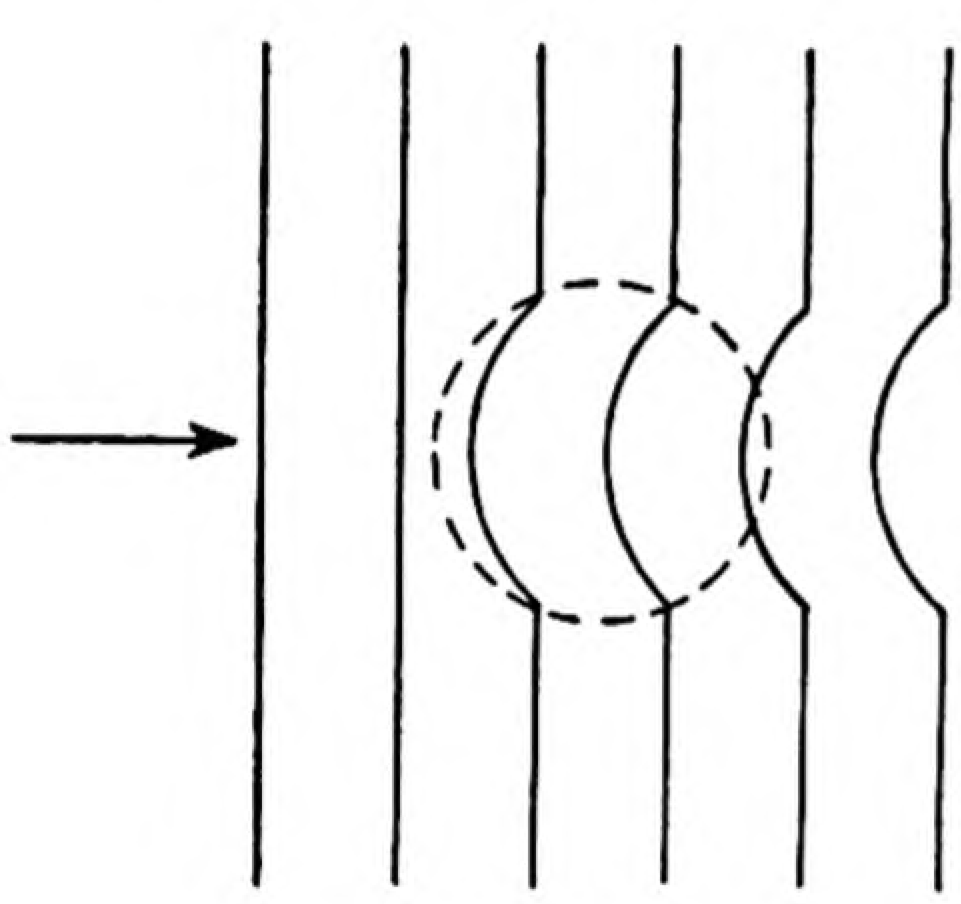


FIG. 2

The wavefront for particles which traverse the sphere lag behind those which do not. The entire wavefront then develops a hemispheroidal depression as it passes through the sphere. By a direct application of Huygens' and Babinet's principles to the further development of the wavefront, van de Hulst] shows that if the scattered wave function is written as $\psi_s(\theta) = R^{-1}f(\theta) \exp(ik_0R)$, then

$$\text{(18}a\text{)} \quad f(\theta) = -a^2k_0 \int_0^{\pi/2} (1 - e^{-2ik_0(m-1)a\cos\gamma}) J_0(z \sin\gamma) \sin\gamma \cos\gamma \, d\gamma,$$

where $u = a \sin\gamma$ and γ is the angle between a chord and its associates normal to the sphere. The parameter z is defined by $z = bk_0 \sin\theta$.

It is interesting that this expression, which was derived by an application of Huygens' principle, is equivalent to the limiting result as $m \to 1$ that follows from the exact formula (10b) by applying the Bessel-function approximations (12a) and (12b) to *both* the numerator and denominator of (10b). On this basis van de Hulst showed that

$$\text{(18}b\text{)} \quad f(\theta) = -a^2ik_0 \int_0^{\pi/2} (1 - e^{-2iak_0(m-1)\cos\gamma}) J_0(z \sin\gamma) \sin\gamma \cos\gamma \, d\gamma,$$

where now

$$z = bk_0\theta$$

The difference of the phase factors of the two forms of $f(\theta)$ comes from a difference in definition of phase in the two analyses. In the limit as $m \to 1$ (18a) yields an expression for the differential-scattering cross section which is equivalent to the corresponding limit of (17).

Equation (18a) can be written in a slightly different form if we choose the distance u of the chord from the center of the sphere as the variable of integration and express the integrand in terms of $x(u)$, the phase changes at n. Then

$$f(\theta) = \frac{k_0}{i} \int_0^\infty (1 - e^{-i\chi(u)}) J_0(k_0 u\theta) u \, du. \tag{18b}$$

4. The Approximating Sphere. Let us now consider a class of spherically symmetrical soft scatterers with the properties:

$$\text{i. } k^2(r) - k_0^2 = A^2 g\left(\frac{r}{b}\right). \tag{19}$$

ii. $|g(x)|$ is a monotone nonincreasing function of x which approaches 0 sufficiently rapidly as $x \to \infty$ so that $\int_0^\infty xg(x)\,dx$ and $\int_0^\infty x^2 g^2(x)\,dx$ exist. The sign of $g(x)$ is postulated to be the same for all values of x.

iii. A and b are constants which measure the strength and range of scatterer. A is small in soft scatterers.

In the case of a homogeneous spherical obstacle of radius b and propagation constant k_1,

$$g\left(\frac{r}{b}\right) = \begin{cases} 1, & (r < b), \\ 0, & (r > b), \end{cases} \qquad A^2 = k_1^2 - k_0^2. \tag{20}$$

The scattered wave function is given in (7) as a volume integral over the internal wave function. Hence if there is a scatterer whose $g(x)$ approximates the required $g(x)$ and whose internal wave function is known, then the internal wave function which corresponds to the approximate $g(x)$ can be used in (7) to compute an approximate scattered wave function for the $g(x)$ of interest.

Let $B^2 g_0(r/a)$ be the function $k_0^2(r) - k_0^2$ which is associated with the approximating scatterer. Then for a given scatterer of strength A and range b there is an approximating scatterer with $k^2(r) - k_0^2 = B^2 g_0(r/a)$ which is most closely related to it in the least square sense so that [14]

$$I = \int \left[A^2 g\left(\frac{r}{b}\right) - B^2 g\left(\frac{r}{a}\right)\right]^2 |\psi_0(r)|^2 \, d\tau = \text{minimum}. \tag{21}$$

In the approximation method discussed in this and the next section we shall find scattered wave functions from (6) by using the internal wave function of an approximating scatterer whose parameters are chosen so that (21) is satisfied.

The simplest approximating function $g_0(x)$ that one might choose when $|g(x)|$ is nonincreasing is the step function:

(22)
$$B^2 g_0\left(\frac{r}{a}\right) = \begin{cases} k_1^2 - k_0^2, & (\text{if } r < a), \\ 0, & (\text{if } r > a). \end{cases}$$

Then

(23a)
$$g_0(x) = \begin{cases} 1, & (x < 1), \\ 0, & (x > 1), \end{cases}$$

and

(23b)
$$B^2 = k_1^2 - k_0^2 \begin{cases} > 0 \text{ if } g\left(\frac{x}{b}\right) > 0 \text{ for all } x, \\ < 0 \text{ if } g\left(\frac{x}{b}\right) < 0 \text{ for all } x. \end{cases}$$

This approximation corresponds to computing the internal wave function by replacing the original scatterer by a homogeneous sphere. Since internal and scattered wave functions of a soft spherical scatterer are known [see equation (14b)], one can substitute their analytical expression and (22) and (21) to obtain the values of k_1 and a which minimize (21).

The wave function $\psi_0(r)$ which corresponds to $B^2 g_0(x/a)$ [equation (22)] is [see equation (14b)], to the first order in $k_1 - k_0$,

$$\begin{aligned} \psi_0(r) &= \frac{2k_1}{k_1 + k_0} e^{ia(k_1 - k_0)} \exp\,(ik_1 \mathbf{r} \cdot \mathbf{s}_0), & (|r| < a), \\ &= \exp\,(ik_0 \mathbf{r} \cdot \mathbf{s}_1) + \psi_s(r), & (|r| > a), \end{aligned}$$

where $\psi_s(r)$ is the scattered wave function. It is proportional to $k_1 - k_0$ and, for large r, varies as $1/r$. Hence, for soft scatterers we can neglect $\psi_s(r)$ as compared with $\exp ik_0(\mathbf{r} \cdot \mathbf{s}_0)$. Physically this corresponds to using the sphere internal wave function inside the approximating sphere and the Born approximation for the "tail" of the scatterer, which lies outside the sphere. Hence we write

(24)
$$\psi_0(r) = \begin{cases} \dfrac{2k_1}{k_1 + k_0} e^{ia(k_1 - k_0)} \exp\,(ik_1 \mathbf{r} \cdot \mathbf{s}_1), & (|r| < a), \\ \exp ik_0(\mathbf{r} \cdot \mathbf{s}_0), & (|r| > a). \end{cases}$$

Equation (21), which defines a and k_1, becomes

(25)
$$I = \frac{4k_1^2}{(k_1 + k_0)^2} \int_0^a \left[(k_1^2 - k_0^2) - A^2 f\left(\frac{r}{b}\right) \right]^2 r^2\,dr + A^4 \int_a^\infty f^2\left(\frac{r}{b}\right) r^2\,dr.$$

The equation $\partial I/\partial a = 0$ and $\partial I/\partial k_1$, which determine the values of a and k_1 which minimize (25), are

(26a)
$$A^2 f\left(\frac{a}{b}\right) = \frac{2k_1(k_1^2 - k_0^2)}{3k_1 + k_0}$$

and

$$(26b)\quad 2k_1^2(k_1+k_0)\int_0^a\left[(k_1^2-k_0^2)-A^2g\left(\frac{r}{b}\right)\right]r^2\,dr + k_1\int_0^a\left[(k_1^2-k_0^2)-A^2g\left(\frac{r}{b}\right)\right]^2 r^2\,dr = 0.$$

In the limit as $k_1 - k_0 \to 0$ and $A^2 \to 0$, the second integral in (26b) becomes negligible, and (26b) reduces to

$$(27a)\qquad \frac{2}{3}\,g(\alpha)a^3 = \int_0^\alpha x^2 g(x)\,dx,$$

where we define

$$(27b)\qquad \alpha = \frac{a}{b}.$$

These equations relate the radius a of the approximating sphere to the range of the scatterer. In the limit, (26a) (which determines k_1) is

$$(28)\qquad k_1^2 - k_0^2 \simeq 2A^2f(\alpha).$$

5. Approximate Differential-scattering Cross Sections of Spherically Symmetrical Scatterers. Once the parameters a and k_1 of the approximating sphere have been obtained, the approximate scattered wave function for a spherically symmetrical scatterer (with the properties i, ii, and iii listed in the beginning of the last section) can be calculated by substituting $\psi(r)$ [equation (24)] into (6).

The fundamental equation then becomes

$$(29)\qquad \psi_s(R) \sim \frac{A^2e^{iRk_0}}{4\pi R}\int \psi_0(r)g\left(\frac{r}{b}\right)\exp\left[-ik_0(\mathbf{r}\cdot\mathbf{s}_1)\right]d\tau,$$

where $\psi_0(r)$ is defined by (24). After performing the angle integrations in the same manner as was done in the derivation of (16b) and (17b), we have

$$(30a)\quad \psi_s(R) \sim A^2R^{-1}e^{iRk_0}\left\{\frac{2k_1}{k_1+k_0}e^{i(k_1-k_0)}\int_0^\infty g\left(\frac{r}{b}\right)\frac{r^2\sin r\omega}{r\omega}\,dr + \int_a^\infty r^2 g\left(\frac{r}{b}\right)\left[\frac{\sin ru}{ru} - \nu\frac{\sin r\omega}{r\omega}\right]dr\right\},$$

$$(30b)\qquad \nu = \frac{2k_1}{k_1+k_0}\exp ia(k_1-k_0),$$

$$(30c)\qquad \omega^2 = k_1^2 + k_0^2 - 2k_1k_0\cos\theta,$$

$$(30d)\qquad u = 2k_0\sin\tfrac{1}{2}\theta.$$

The approximate scattered wave function can be determined easily for a screened coulomb scatterer. In that case

$$g(x) = x^{-1}e^{-x},$$

and after integration

$$(31)\quad \psi_s \simeq A^2 e^{ik_0R}R^{-1}\left[\frac{2k_1}{k_1+k_0}\,e^{ia(k_1-k_0)}\left(\frac{b^3}{1+\omega^2b^2} - \frac{b^2e^{-a/b}}{\omega}\,\frac{\sin\omega + b\omega\cos\omega a}{1+b^2\omega^2} + \frac{b^2}{u}\,e^{-a/b}\,\frac{\sin ua + bu\cos ua}{1+b^2u^2}\right)\right].$$

The differential-scattering cross section is

$$(32)\qquad \sigma_s(\theta) = R^2|\psi_s|^2.$$

In the first approximation it can be shown [14] that

$$\alpha = \frac{a}{b} = 0.807,$$
$$A^2 = \pm 0.379(k_1^2 - k_0^2).$$

The corresponding results can be obtained with a Gaussian scatterer

$$g(x) = \exp(-x^2).$$

Then

$$(33)\quad \psi_s \simeq \frac{A^2e^{iRk_0}}{R}\left\{\frac{2k_1}{k_1+k_0}\,\frac{b^3\pi^{\frac{1}{2}}}{4}\,e^{i(k_0-k_1)}e^{-\omega^2b^2/2} + \frac{b^3}{2}\,e^{-\alpha^2}[I(u) - \nu I(\omega)]\right\},$$

where

$$I(y) = \alpha\,\frac{\sin ya}{ya} + e^{+\alpha^2}\int_\alpha^\infty e^{-x^2}\cos ybx\,dx$$

and the differential-scattering cross section is given by (32). Here in the first approximation [14]

$$\alpha = \frac{a}{b} = 1.235,$$
$$\pm A^2 = 1.744(k_1^2 - k_0^2).$$

The variables ν, u, and ω are defined by (30). As before, the differential-scattering cross section is given by (32).

These expressions differ somewhat from those given by the authors in another paper [14]. In that paper it was assumed that

$$\psi_0(r) = \frac{2k_1}{k_0+k_1}\exp\left[ia(k_1-k_0) + ik_1(\mathbf{r}\cdot\mathbf{s}_0)\right], \qquad \text{(for all } r\text{)}.$$

The use of this approximation was based on the fact that $\psi_0(r)$ and $\partial\psi_0(r)/\partial r$ are both continuous at $r = a$ and $k^2(r) - k_0^2 \to 0$ rapidly as $r \to \infty$.

An error can be expected when the "tail" of the scatterer outside the approximating sphere extends out a large distance. One way of observing the existence of the error is to calculate the total-scattering cross section of the scatter as its range becomes large. In the case of the Gaussian scatterer the ratio of the total-scattering cross section to the "geometrical" cross section

πb^2 approaches zero as $b \to \infty$ when one applies the formulas of the earlier paper [14]. This means that the tail is not given sufficient weight, for as long as a scatterer does not have a sharp boundary, this ratio should become infinite (but not as rapidly as it would in the Born approximation). We have given the tail more weight in the calculation made above by using the Born approximation in the scattering by the tail. Although as the scatterer becomes large this overemphasizes the tail, formula (30*a*) should be valid over a considerably wider range than either the Born approximation or the simplified equivalent-sphere approximation used in the earlier paper [14].

Another approach to the problem of scattering by a soft obstacle has been made by Glauber [15]. He, independently of van de Hulst, has shown that equation (18*b*) is valid for a wide variety of spherically symmetrical scatterers in addition to the homogeneous sphere. Glauber has derived (18*b*) for the general case in several different ways. One is to replace the integral equation (6) by an approximate differential equation which is valid for very soft scatterers. The internal wave function is found by solving the approximate differential equation. The scattered wave function is derived by substituting the approximate internal wave function into (6) and integrating. Also, by using the expansion of the scattered wave function in partial waves he improved (18*b*) by replacing the term $z = a\theta k_0$ by $z = 2ak_0 \sin \frac{1}{2}\theta$.

6. Total-scattering Cross Section for Very Soft Scatterers. The total-scattering cross section σ can be obtained by integrating $\sigma_s(\theta)$ over all angles,

$$\sigma = 2\pi \int_0^\pi \sigma_s(\theta) \sin \theta \, d\theta. \tag{34}$$

It is also related to the scattered wave function

$$\psi_s(\theta) = \frac{e^{ik_0R}}{R} f(\theta) \tag{34a}$$

through the formula

$$\sigma = \frac{4\pi}{k_0} \operatorname{Im} f(\theta). \tag{35}$$

This equation is equivalent to (34) only when (34*a*) is exact. In some approximations (for example, in the Born approximation) the approximate $f(\theta)$ is especially poor in the region of $\theta = 0$. Hence, even though it is much more convenient to use (35), the errors in the total-scattering cross section can better be identified with those of the $f(\theta)$ as a function of θ if (34) is used. Actually an alternative method of selecting parameters of approximating potentials is to make (34) and (35) and other similar formulas yield equivalent results. In the previous papers of this series total-scattering cross sections were derived from (34).

Recently Glauber [15] developed a new formula for the total-scattering cross section of very soft scatterers which can be deduced without a detailed knowledge of the scattered wave function. It is a generalization of a result obtained

by van de Hulst [13] from (18*b*) and (35). We shall give a derivation of his formula and discuss several special cases using it.

Let us again apply the arguments of van de Hulst (see end of Sec. 3) to the scattering of a particle by a square-well (or step) potential. Our formula (14*b*) for the internal wave function of a very soft sphere (as $k \to 0$) shows that a particle which travels along a chord R is not deflected from its original path (as $k_1 \to k_0$) but that its wavelength is changed while it is in the range of the scatterer. The total change in phase by passage through the field depends on the distance u of closest approach of the particle to the scattering center. We let $\chi(u)$ be this total change in phase and $\exp(izk_0)$ be the wave function of a free particle in the absence of the scatterer. Then at a large distance from the scattering center, a particle which passes along a chord R has the wave function

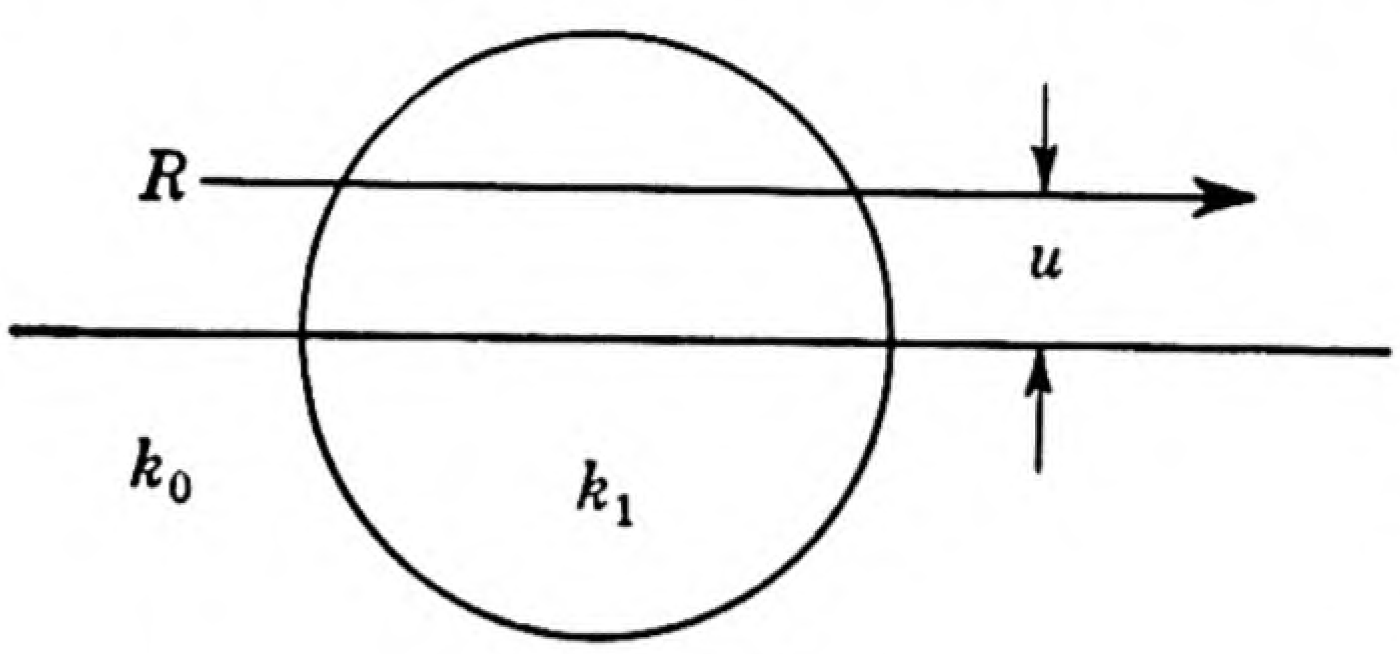


FIG. 3

$$\exp[izk_0 + i\chi(u)].$$

The variation of the wave function due to scattering is

$$e^{izk_0+i\chi(u)} - e^{izk_0}$$

so that the amplitude of the scattered wave is

$$|e^{izk_0+i\chi(u)} - e^{izk_0}|^2 = 4\sin^2 \tfrac{1}{2}\chi(u). \tag{36}$$

This result seems to be quite general. In the limit as the strength of a spherically symmetrical scatterer diminishes, there is no refraction but only a change in phase of the incident particle. Hence, in the limit the total-scattering cross section of a vanishingly weak spherically symmetrical scatterer becomes

$$\sigma = 2\pi \int_0^\infty 4\sin^2 \frac{\chi(u)}{2}\, u\, du. \tag{37}$$

This expression, originally derived by Glauber, follows directly from the Glauber–van de Hulst formula by combining (35) with (18*b*).

The calculation of $\chi(u)$ is easily performed when

$$k^2(r) - k_0^2 = A^2 g\left(\frac{r}{b}\right) \tag{38}$$

so that $k(r) - k_0 \sim \frac{1}{2}(A^2/k_0)g(r/b)$ as $A \to 0$. In the limit as $A \to 0$, a useful parameter is

$$y = \left(\frac{A}{k_0}\right)^2 bk_0. \tag{39}$$

In the special case (38), (37) reduces to (after letting $x = u/b$)

$$\frac{\sigma}{\pi b^2} = 8 \int_0^\infty x \sin^2 \frac{\chi(bx)}{2}\, dx. \tag{40}$$

We find $\chi(bx)$ by noting that the change in phase per unit length at z along the chord R is

$$[k(\sqrt{u^2 + z^2}) - k_0]$$

so that when (38) is satisfied

$$\chi(u) = \frac{y}{2b} \int_{-\infty}^{\infty} g\left(\sqrt{\frac{u^2 + z^2}{b^2}}\right) dz.$$

Hence $\sigma/\pi b^2$ is given by (40) when

$$\chi(bx) = \frac{y}{2} \int_{-\infty}^{\infty} g(\sqrt{x^2 + v^2})\, dv, \tag{41}$$

and y is defined by (39).

In some special cases

(a) $$g(x) = \begin{cases} \dfrac{k_1^2 - k_0^2}{A^2} = 1, & (x < 1), \\ 0, & (x > 1), \end{cases}$$

$$y = 2b(k_1 - k_0).$$

Then

$$\chi(bx) = \begin{cases} y(1 - x^2)^{\frac{1}{2}}, & (\text{if } x < 1), \\ 0, & (\text{if } x > 1). \end{cases}$$

(b) $g(x) = e^{-x^2}$ yields $\chi(bx) = \frac{1}{2} y \pi^{\frac{1}{2}} e^{-x^2}$.

(c) $g(x) = x^{-3}$ yields $\chi(bx) = \dfrac{y}{x^2}\cdot$

(d) $g(x) = e^{-x}$ yields $\chi(bx) = yK_1(x)$.

(e) $g(x) = x^{-1}e^{-x}$ yields $\chi(bx) = yK_o(x)$.

The function $K_j(x)$ is the jth Bessel function of the second kind of purely imaginary argument (see Watson [16, p. 78]).

For a uniform spherical scatterer

$$\frac{\sigma}{\pi b^2} = 8 \int_0^1 x \sin^2 \left[\frac{y}{2}(1 - x^2)^{\frac{1}{2}}\right] dx = 4 \left(\frac{\pi}{2y}\right)^{\frac{1}{2}} H_{\frac{3}{2}}(y), \tag{42}$$

where $H_{\frac{3}{2}}(y)$ is the Struve function [16] of order $\frac{3}{2}$ defined by

$$H_{\frac{3}{2}}(y) = \left(\frac{y}{2\pi}\right)^{\frac{1}{2}} (1 + 2y^{-2}) - \left(\frac{2}{\pi y}\right)^{\frac{1}{2}} (\sin y + y^{-1} \cos y)$$

and $y = 2ak_0(m - 1)$ with $m = k_1/k_0$. This is exactly the result obtained by van de Hulst [13] and Hart and Montroll [12]. As the range of the scatterer approaches ∞ and hence as $y \to \infty$, $\sigma/\pi b^2 \to 2$.

In the case of the Gaussian scatterer

$$(43)\qquad \begin{aligned}\frac{\sigma}{\pi b^2} &= 8\int_0^\infty x\sin^2\left(\frac{1}{4}y\pi^{\frac{1}{2}}e^{-x^2}\right)dx = 2\int_0^{\frac{1}{4}y\pi^{\frac{1}{2}}}\frac{(1-\cos t)}{t}\,dt\\ &= 2\log\frac{\gamma y\pi^{\frac{1}{2}}}{2} - 2\,\mathrm{Ci}\,\frac{y\pi^{\frac{1}{2}}}{2},\end{aligned}$$

where $\log\gamma = 0.5772$, and

$$\mathrm{Ci}\,x = \log\gamma x - \int_0^x \frac{1-\cos t}{t}\,dt.$$

The function $\mathrm{Ci}\,x$ is tabulated in Janke-Emde. As $x\to\infty$, $\mathrm{Ci}\,x\to 0$. Equation (43) was first derived by Glauber [15].

The Born approximation of the total-scattering cross section is obtained from (40) by using the hypothesis that the phase change resulting from the passage of the incident wave through the scatterer is very small so that $\sin\frac{1}{2}\chi(bx)\simeq\frac{1}{2}\chi(bx)$. Then

$$(44)\qquad \frac{\sigma}{\pi b^2}\sim\frac{y^2}{2}\int_0^\infty x\left[\int_{-\infty}^{\infty} g(\sqrt{(x^2+v^2)}\,dv\right]^2 dx.$$

Since y is proportional to b, the range of the scatterer, the Born approximation implies that as $b\to\infty$, the total-scattering cross section increases with the fourth power of the range. This is incorrect for all scatterers considered. On the other hand, as $b\to 0$, the Born-approximation expression is in agreement with the results of more exact analysis.

There is a case of a highly singular potential, $g(x) = x^{-3}$, for which the Born approximation (44) diverges and, even as $b\to 0$, σ does not approach b^4. In this case $\chi(bx) = yx^{-2}$ so that a formal application of (40) yields

$$(45)\qquad \begin{aligned}\frac{\sigma}{\pi b^2} &= 8\int_0^\infty x\sin^2\frac{1}{2}\frac{y}{x^2}\,dx\\ &= 2y\int_0^\infty u^{-2}\sin^2 u\,du = \pi y.\end{aligned}$$

There is some question as to whether equation (40) is valid for such a singular potential. Hence, equation (45) should be checked by a more rigorous analysis.

The equation analogous to (40) for the total-scattering cross section per unit length of an infinite, soft, circularly symmetrical cylinder is

$$(46)\qquad \frac{\sigma}{2b} = 2\int_{-\infty}^\infty \sin^2\frac{\chi(bx)}{2}\,dx.$$

As an example of an application of (46) let us consider the case of a uniform homogeneous circular cylinder of radius b. Then $\chi(bx)$ is given by the case (a) under equation (41). Hence

$$\frac{\sigma}{2b} = 4\int_0^1 \sin^2\left[\frac{y}{2}(1-x^2)^{\frac{1}{2}}\right]dx.$$

After letting $x = \sin\theta$, and integrating by parts, we have

$$\frac{\sigma}{2b} = 2y \int_0^{\pi/2} \sin^2\theta \sin(y\cos\theta)\, d\theta = \pi H_1(y),$$

where $H_1(y)$ is the Struve function of order 1. This is a result previously obtained by van de Hulst [17] and Montroll and Hart [12].

7. Scattering by Systems Which Contain Several Scatterers. So far we have discussed only the scattering by isolated obstacles. In many physical applications of scattering techniques the influence of the scattering pattern of one scatterer on that of another becomes important. This influence can occur in several ways. The path length from source to observer may be different in two of the scatterers (see Fig. 4a) so that a phase difference occurs between the two scattered waves and interference can take place. Also, multiple scattering can occur (see Fig. 4b). The effect of multiple scattering increases with the concentration of scatterers. Here we shall be concerned

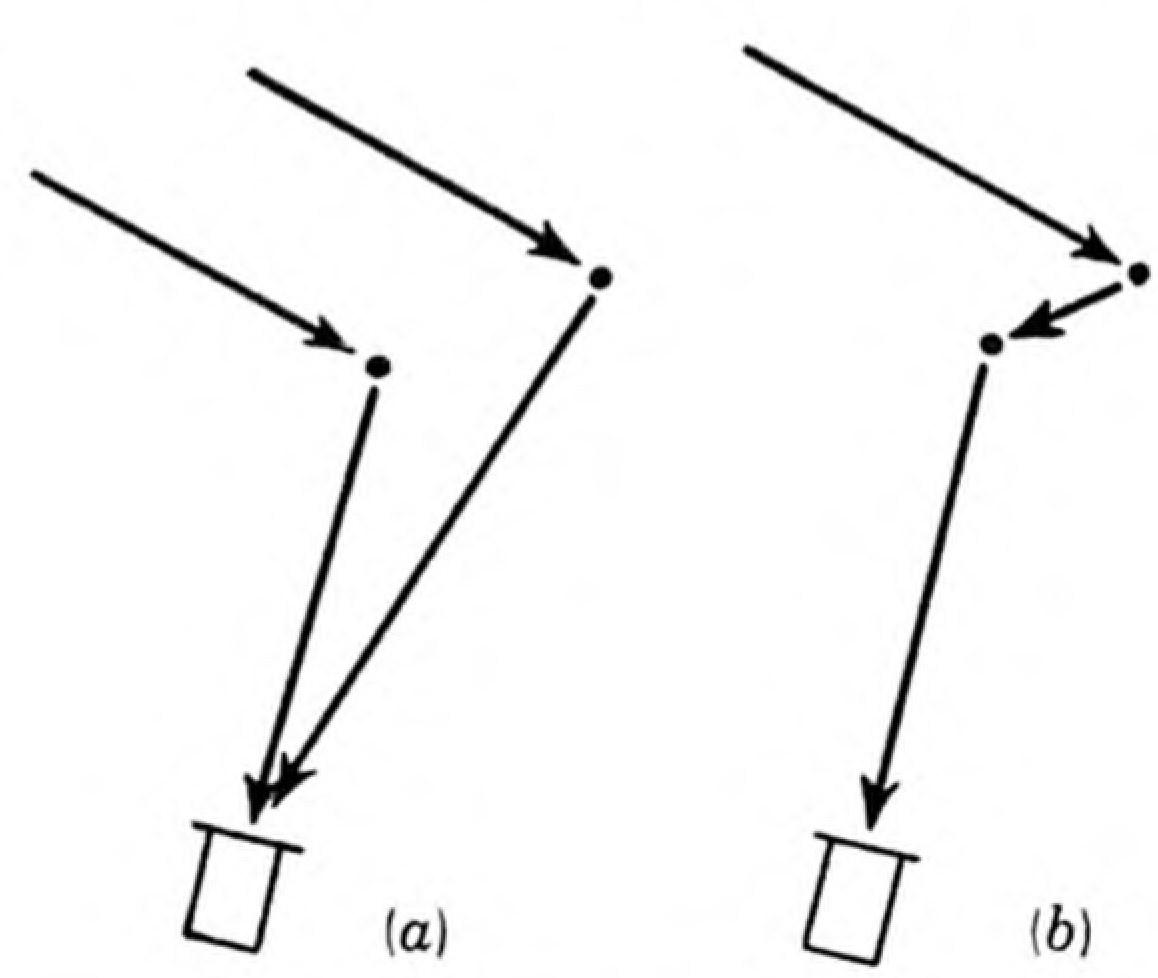


FIG. 4

only with the former effect. We shall assume that the concentration of scatterers is so small that the multiple scattering is negligible.

When the range of at least one of several different species in the same system is somewhat larger than that of the incident wave, the difference in the effective path lengths of waves which pass through various species also contributes to the interference effect. This phenomenon is exaggerated as the difference in the average relative index of refraction of the various species increases. We shall now discuss the theory of this situation. The importance of the effect was first noted by Shomaker and Glauber [18] in their investigation of the theory of scattering of electrons by molecules containing heavy atoms.

Consider a fixed distribution of atoms of a molecule in gas, liquid, or powder which are located at points $\{v_j\}$ with respect to some preassigned origin. The scattered wave due to the jth particle is then, at a point $\mathbf{r}(|\mathbf{r}| \gg |\mathbf{r}_j|)$, given by

$$\psi_j = r^{-1} \exp(irk_0) f_j(\theta) \exp[ik_0(\mathbf{s}_0 - \mathbf{s}_1) \cdot \mathbf{r}_j]. \tag{47}$$

Here $\mathbf{s}_0$ and $\mathbf{s}_j$ are, respectively, unit vectors in the direction of the incident and scattered radiation, $\cos\theta = \mathbf{s}_0 \cdot \mathbf{s}_1$, and $f_j(\theta)$ is the function defined in (34a).

The sum of the individual scattered wave functions is the total scattered wave function

$$\psi = \sum \psi_j = \frac{e^{irk_0}}{r} \sum \exp\,[ik_0(\mathbf{s}_0 - \mathbf{s}_1) \cdot \mathbf{r}_j] f_j(\theta). \tag{48}$$

In gases, normal liquids, and powders the molecules are randomly oriented. Hence the total scattered intensity is obtained by averaging the intensity for a single molecule over all possible orientations. Then

$$I(\theta) = r^2 \langle |\psi|^2 \rangle = \sum_{i,j} f_i f_j^* \frac{\sin sr_{ij}}{sr_{ij}}, \tag{49}$$

where $s = 2k_0 \sin \frac{1}{2}\theta$.

$r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between ith and jth atoms.

When the scattering amplitudes are taken to be real (as they are in the Born approximation), the only interference effect is that due to the special distribution of the atoms in a molecule. However, if the internal phase shifts differ for different types of atoms, the interferences between waves scattered from these different atoms will be modified.

The variation of the intensity of scattered radiation with angle in a single crystal is also obtained from (48). However, the simplification of random angular orientation which leads to (49) cannot be used.

If we let

$$f_j = \rho_j \exp i\varphi_j, \qquad \delta_{jk} = \varphi_j - \varphi_k, \tag{50a}$$

then

$$I(\theta) = \sum \rho_j^2 + 2 \sum_{j>k} \rho_j \rho_k \cos \delta_{jk} \frac{\sin sr_{jk}}{sr_{jk}}. \tag{50b}$$

As a special application of this formula let us consider the scattering of electrons by a molecule of the form AB_n, where the n B atoms are symmetrically spaced about the A atom. Let $r_{11}, r'_{11}, r''_{11}, \cdots$ represent the distances between the (B,B)-pairs and $A_{11}, A'_{11}, A''_{11}, \cdots$ the number of pairs at those distances. Also, let r_{12} be the distance between an (A,B)-pair. Then

$$I(\theta) - \rho_2^2 - n\rho_1^2 = 2\rho_1\rho_2 \cos \delta_{12} \frac{\sin sr_{12}}{sr_{12}} + \rho_1^2 \left(A_{11} \frac{\sin sr_{11}}{sr_{11}} + A'_{11} \frac{\sin sr'_{11}}{sr'_{11}} + \cdots \right). \tag{51}$$

The combination $\cos \delta_{12} \sin sr_{12}$ might be written $\frac{1}{2}[\sin (sr_{12} - \delta_{12}) + \sin (sr_{12} + \delta_{12})]$.

The procedure used to determine molecular structures by electron diffraction is to guess a set of reasonable geometrical structures and sets of interatomic distances (the numbers $r_{11}, r'_{11}, r''_{11}, r_{12}, \cdots$ in the case at hand) and calculate a theoretical $I(\theta)$ for each of these models. That structure whose calculated $I(\theta)$ has maxima and minima at points closest to the observed ones

is considered to be the correct one. If the theoretical calculation is made using the Born approximation (as is usually done), the phase differences δ_{ij} are zero so that the locations of maxima and minima determined on that basis might not be the same as those given by (51). It can be shown that in molecules such as UF_6 which contain atoms of widely different atomic numbers, the phase shifts δ_{ij} may be sufficiently large even to change the sign of the first term on the right-hand side of equation (51).

Let us choose the screened coulomb potential

$$V(r) = Ze^2r^{-1}\exp\left(\frac{-r}{b}\right), \qquad (b = a_0Z^{-\frac{1}{3}})$$

as the scattering potential of an atom of atomic number Z. The constant a_0 is the Bohr radius; $a_0 = 0.528 \times 10^{-8}$ cm. Although the phase shift δ_{12} is angular-dependent, we can obtain its order of magnitude by considering only the first term inside the large parentheses of (31). In the notation of Sec. 4 this potential is equivalent to

$$g(x) = x^{-1}e^x, \qquad A^2 = -2me^2\frac{Z^{\frac{4}{3}}}{a_0}.$$

For an order of magnitude calculation of δ_{12} we write

$$\delta_{12} \simeq [a(k_1 - k_0)]_A - [a(k_1 - k_0)]_B,$$

where the subscripts A and B refer to atomic species A and B. From the equations below (31) for the equivalent sphere radius a and index k_1 in terms of b and A^2 we can finally show (using the relativistic expression for k_0) that

$$\delta_{12} \simeq \frac{(\lambda_0/a_0)(Z_A - Z_B)}{(E + E^2)^{\frac{1}{2}}} \tag{52}$$

where λ_0/a_0 is the ratio of the Compton wavelength to the Bohr radius of the hydrogen atom. It is to be noted from (52) that as the energy of the incident electron increases, $\delta_{12} \rightarrow 0$, the Born-approximation value. As the difference in the atomic numbers increases, the phase δ_{12} increases. Hence at very high energies and for atoms whose atomic numbers are close to each other, one is justified in using the Born approximation in electron-diffraction analysis.

In Table I we have calculated approximate values of δ_{12} and $\cos\delta_{12}$ for several molecules of the form AB_n. It should be remembered that δ_{12} is actually a function of angle, so that the values given are rough averages. For

TABLE I

Molecule	δ_{12}	$\cos\delta_{12}$
CCl_4	0.4	0.9
SiF_4	0.2	1.0
UF_6	3.0	−1.0
WF_6	2.5	−0.8
MoF_6	1.3	0.3

those values of cos δ_{12} close to 1 the Born approximation is quite valid. On the other hand in the case of UF_6, WF_6, and MoF_6, the peaks and minima are shifted considerably from the Born-approximation values, so much so that for many years the wrong structures were assigned to these molecules in order to make the theoretical and experimental scattering patterns fit. The importance of the phase shift is the analysis of experimental electron-diffraction data, as was first pointed out by Shomaker and Glauber. The calculations in Table I were made for 40-kev electrons.

It can be shown that to a first approximation the effect of the internal phase shifts on crystals is to change the relative intensities but not the positions of the diffraction maxima.

The authors are indebted to Dr. R. Glauber for several informative discussions and for communicating some of his results to them before publication. Some of Glauber's results have also been derived by Moliere [19].

Bibliography

1. N. F. Mott and H. S. W. Massey, The theory of atomic collisions, Oxford University Press, New York, 1949, p. 114.
2. Lord Rayleigh, Proc. Roy. Soc. London Ser. A vol. 84 (1910) p. 25; vol. 90 (1914) p. 219.
3. R. Gans, Ann. Phys. vol. 76 (1925) p. 29.
4. M. Born, Zeit. Phys. vol. 37 (1926) p. 863; vol. 38 (1926) p. 803.
5. Lord Rayleigh, *The theory of sound*, Vol. 2, Dover Publications, New York, reprint ed., 1945, p. 272.
6. G. Mil, Ann. Phys. vol. 25 (1908) p. 37.
7. R. Hart and E. Montroll, J. Appl. Phys. vol. 22 (1951) p. 376.
8. P. Debye, Ann. Phys. vol. 30 (1909) p. 57.
9. *Tables of scattering functions for spherical particles*, Nat. Bur. Standards Appl. Math. Ser. 4, Washington (1948).
10. R. O. Gumprecht and C. M. Sliepcevich, *Tables of Light-scattering functions for spherical particles*, University of Michigan Press, Ann Arbor, 1951.
11. P. M. Morse, *Vibration and sound*, McGraw-Hill Book Company, Inc., New York, 1st ed., 1936.
12. E. Montroll and R. Hart, J. Appl. Phys. vol. 22 (1951) p. 1278.
13. H. C. van de Hulst, Recherches astron. Observatoire d'Utrecht vol. 11 (1946).
14. E. Montroll and J. M. Greenberg, Phys. Rev. vol. 86 (1952) p. 889.
15. R. Glauber, Private communication.
16. G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge University Press, London, 2d ed., 1948.
17. H. C. van de Hulst, Astrophys. J. vol. 112 (1950) p. 1.
18. V. Shomaker and R. Glauber, Nature vol. 170 (1952) p. 290.
19. G. Moliere, Zeit. Naturforsch. vol. 2 (1947) p. 133.

Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Md.
Rensselaer Polytechnic Institute, Troy, N.Y.

WAVE PROPAGATION IN HELICAL COMPRESSION SPRINGS[1]

BY

E. H. LEE

Abstract. In the past, spring surges have been analyzed on the assumption that adjacent coils do not collide during the motion. In many cases of compression springs subjected to impact this condition is violated. In this paper a theory of spring surges including coil closure is formulated, both for inelastic and elastic coil-impact conditions. Problems of the former type closely resemble certain problems in the propagation of plastic waves in compression, and the mathematical techniques developed for that work can be utilized. Perfectly elastic coil-on-coil impact demands a quite different type of mathe-

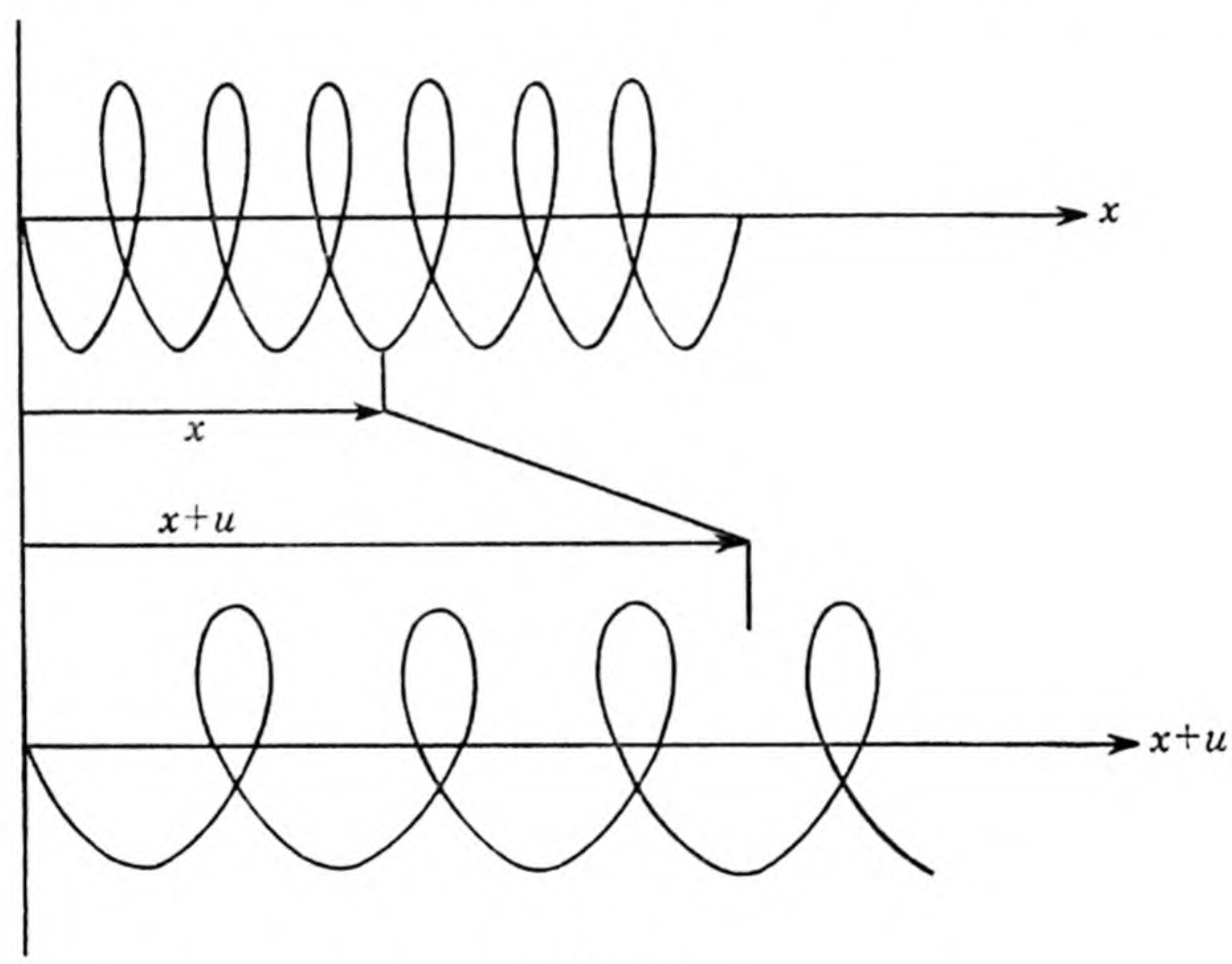


FIG. 1

matical analysis associated with the propagation of force doublets along the elastic spring, resulting in intermittent pulses of compression. Some simple boundary-value problems are analyzed, with an indication of the generalization to more complex problems. The accuracy to be expected from the type of analysis suggested is discussed.

1. Introduction. We are concerned with the motion of a helical spring as depicted diagrammatically in Fig. 1. In practice such springs are commonly designed on the assumption of uniform force and strain distribution throughout the length of the spring, with a resulting linear relation between the transmitted force and the relative displacement of the ends. However, under conditions of dynamic loading, inertia forces of the spring itself modify the force transmitted, and wave-propagation effects occur. Such waves traveling through the

[1] The results presented in this paper were obtained in the course of research sponsored by the Ballistics Research Laboratory, Aberdeen Proving Ground, under Contract No. Da-19-020-ORD-798 with Brown University.

spring are known as spring surges, and they have been treated in the literature on the basis of the linear wave equation with constant velocity of propagation (see, for example, Dejuhasz [1]). The development of such a theory is briefly described below. In the case of helical springs loaded in compression, the applied force tends to close the coils together, and if adjacent coils come in contact, virtually no further compressive strain can take place at that section. Such closure will modify the motion based on the linear wave equation, and it commonly occurs in springs subjected to impact. The present paper describes a theory of spring surges taking into account coil closure. To our knowledge this is the first such treatment, previous work being limited to application of the linear wave equation to determine when coil closure first occurs [2], the subsequent motion not being analyzed.

The usual theory of spring surges is based on the consideration of motion of the spring wire parallel to the axis of the spring only, and the assumption that each element of the spring satisfies the force–longitudinal-strain relation of the whole spring. The complete analysis of the motion of a helix leads to the much more complex theory developed by Love [3], but experiment has shown satisfactory agreement with the simpler theory, as demonstrated, for example, by the work of Weibull [4]. We shall therefore base our development on the simple theory, but some consideration of the influences of the approximations involved is given at the end of this section.

In developing the theory it is convenient to use as independent space variable the position x along the unstressed spring as depicted in Fig. 1. Such a Lagrange-type coordinate is used so that large displacements can be considered without complicating the wave equation which governs the motion. The displacement is $u(x,t)$, so that the position of an element of the deformed spring is given by $x + u$. Let f be the compressive force transmitted across the section x; then the equation of motion of an element dx is

$$f_x = -mu_{tt}, \tag{1}$$

where m is the mass of the spring per unit length and subscripts denote partial derivatives. The nominal compressive strain ϵ of an element of the spring is given by the reduction in length per unit initial length and is represented by $\epsilon = -u_x$. Before closure occurs, there is a linear relation between the force and the strain,

$$f = E\epsilon, \tag{2}$$

where E is a constant depending on the dimensions and material of the spring. Substitution in (1) gives the linear wave equation for u,

$$u_{xx} - \frac{m}{E} u_{tt} = 0, \tag{3}$$

with the constant velocity of wave propagation c_0 given by $c_0^2 = E/m$. Thus, as long as the linear relation (2) applies, spring surges travel with constant

velocity when referred to the undisturbed spring coordinate x, even though the displacements may be large. The extensive theory of the linear wave equation can therefore be utilized. This type of analysis of spring surges is equivalent to the analysis of longitudinal waves in a linear elastic material. f replaces the nominal stress, E the Young's modulus, and m the density. Thus in applying this theory we in effect consider a spring as an equivalent elastic rod capable of withstanding large strains and having a low value of Young's modulus and density, which depend both on the material of the spring wire and on the dimensions of the spring.

In connection with the approximation mentioned above which is involved in this type of theory, that it is based on the longitudinal motion only and the static loading relation (2) for the whole spring, some details of the force transmission should be discussed. In loading a helical spring statically the line of force commonly acts along the spring axis so that a single coil is loaded as shown in Fig. $2a$. For flat coiled springs, which include compression springs as discussed in this paper, it is sufficient to consider a single coil as a plane ring normal to the spring axis as depicted in Fig. 2, and the central loading of Fig. $2a$ then produces a constant torque PR in the spring wire and zero bending moment. The resulting constant twist produces a uniformly distributed displacement along the axis of the spring and thus a uniform strain. The inertia forces of the spring wire, however, act at the spring wire radius R, so that their transmission through the spring must be similar to that of the force P shown in Fig. $2b$. Such a force produces a varying torque and bending moment around each coil. If θ measures the angular position of a section from the point of application of P, the torque is $PR(1 - \cos\theta)$ and the bending moment $PR \sin\theta$. These moment components will produce a different displacement pattern from that associated with axial loading. However, with inertia loading distributed around a coil, these moment variations will be averaged out. The torque $PR(1 - \cos\theta)$ has the average value PR and the bending moment $PR \sin\theta$ the average value zero, so that the situation for axial loading is regained. Thus for a spring with many coils, if the average motion around a coil is of interest, and not the detail propagation of a stress wave within a single coil, the deformation associated with static axial loading can form a basis for the analysis. In many cases this averaging influence is enhanced by the circumstance that the spring slides over a cylindrical former or inside a cylinder, which prevents bending of the spring axis and constrains the motion to be of the type associated with axial loading. Thus one would expect the simple theory to give satisfactory results, excluding the details of the deformation of an individual coil, and in fact it is found experimentally [4] to give good

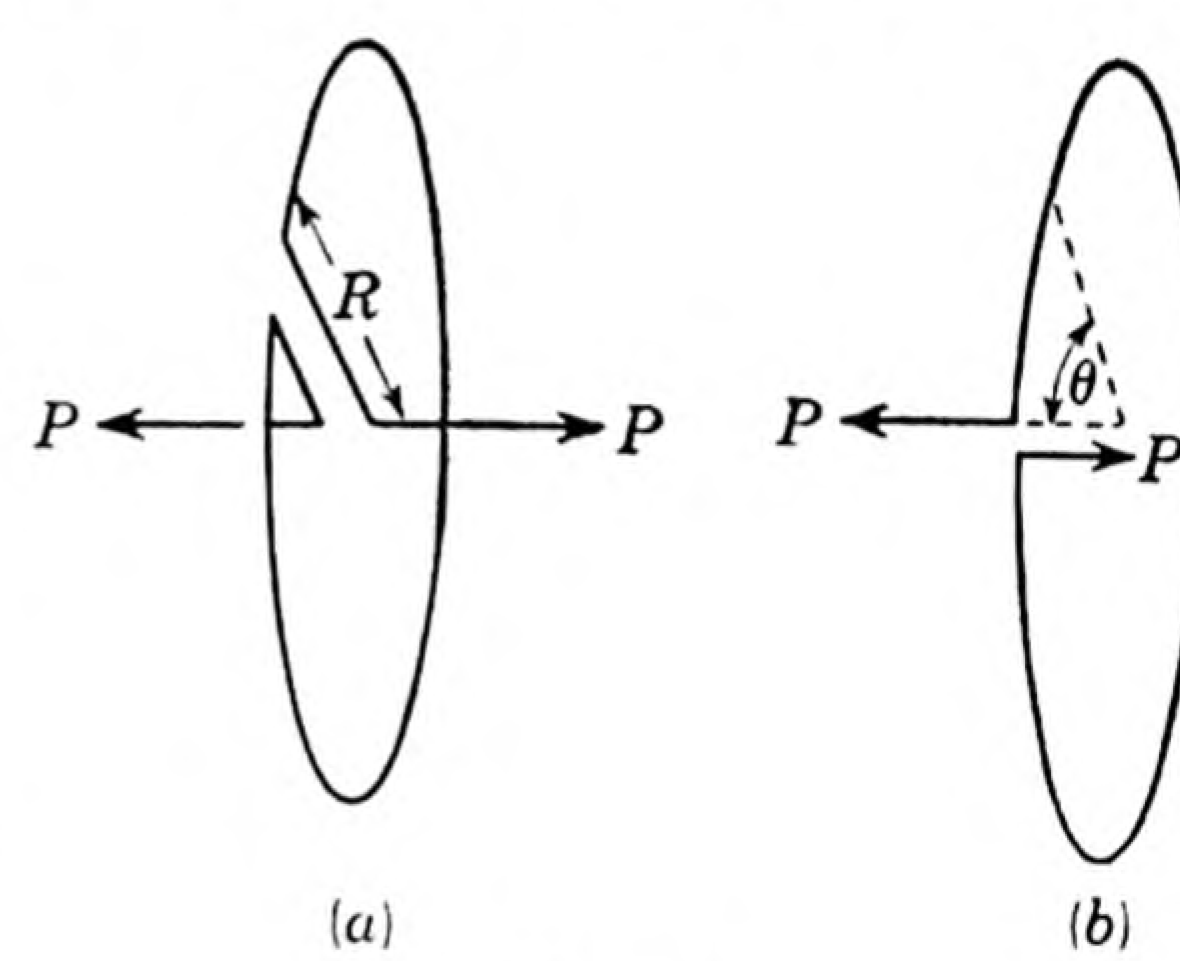


FIG. 2

agreement even in the case of a wavefront of velocity discontinuity in which the averaging process is restricted by the rapid load gradients at the wavefront.

The condition for coil closure is that the approach of sections of adjacent coils with the same angular position is equal to the initial separation of the coils. It is thus associated with the difference in the displacements u at points separated in x by the pitch p of the spring coils and can be written

$$u(x) - u(x + p) = p - d, \tag{4}$$

where d is the axial dimension of the spring wire. Since our theory averages effects around a single coil, it is permissible to replace the finite difference by a derivative and so to consider coil closure to be associated with a fixed upper bound for the compressive strain ϵ, determining the condition

$$\epsilon = -u_x \leq \frac{p - d}{p} = \epsilon_{\max}. \tag{5}$$

Such a maximum-strain condition forms the basis for the coil-closure analysis developed below.

In general when closure occurs, there will be a relative velocity of impact between the adjacent coils, and the subsequent motion will depend on the rebound conditions at the point of contact. We shall adopt the usual theory of contact collision in which the relative velocity of approach and separation after bounce are governed by a coefficient of restitution. In this paper we shall deal with the two extreme cases: inelastic collision and perfect rebound. In the former case, which is dealt with in the next section, contact is maintained after collision until the surfaces are separated by the elasticity of the spring, there being no separating impulse generated by the collision. In the latter case of perfect rebound, the collision impulse which reduced the relative velocity of adjacent coils to zero is repeated to produce bounce with a relative velocity equal to that just prior to contact. As in the usual theory of collision, the relative approach during contact is neglected, for in general it will be small compared with the relative approach prior to closure.

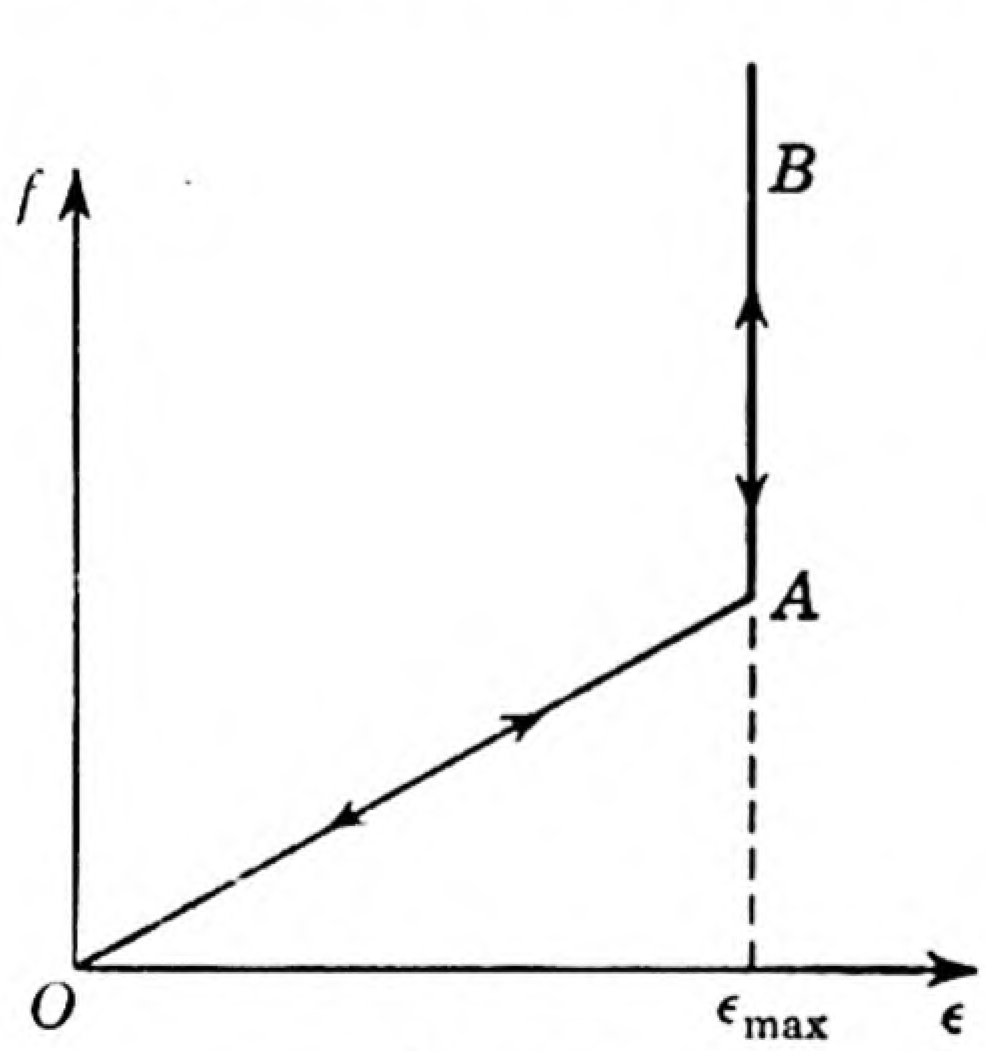


Fig. 3

2. Closure with inelastic coil contact. As mentioned in the previous section, the force-strain relation for this case is that depicted in Fig. 3. The linear relation OA applies prior to closure, and during closure any increase of force can be transmitted without change in strain. If the force transmitted falls below the closure force, the linear relation OA in Fig. 3 again applies. This force-strain relation is a special case of a general functional relation

$$f = f(\epsilon) \tag{6}$$

which replaces the linear relation (2) of the discussion of the previous section. (3) then becomes

$$u_{xx} - \frac{m}{f'} u_{tt} = 0, \tag{7}$$

where f' is the derivative of $f(\epsilon)$. Thus we have the wave equation in which the wave velocity c, given by $c^2 = f'/m$, is a function of the strain, *i.e.*, of $-u_x$. The treatment of boundary-value problems for this equation has been quite fully considered in the analysis of longitudinal plastic waves in rods, as developed by Kármán, Bohnenblust, and Hyers [5]. Thus, in effect, for the case of inelastic closure we can replace the spring by an equivalent nonlinear elastic rod of density m and with the stress-strain relation $f(\epsilon)$ depicted in Fig. 3. The analysis developed in [5] is limited to cases in which the $f(\epsilon)$ relation is concave downward, for which a wavefront of increasing stress spreads out as

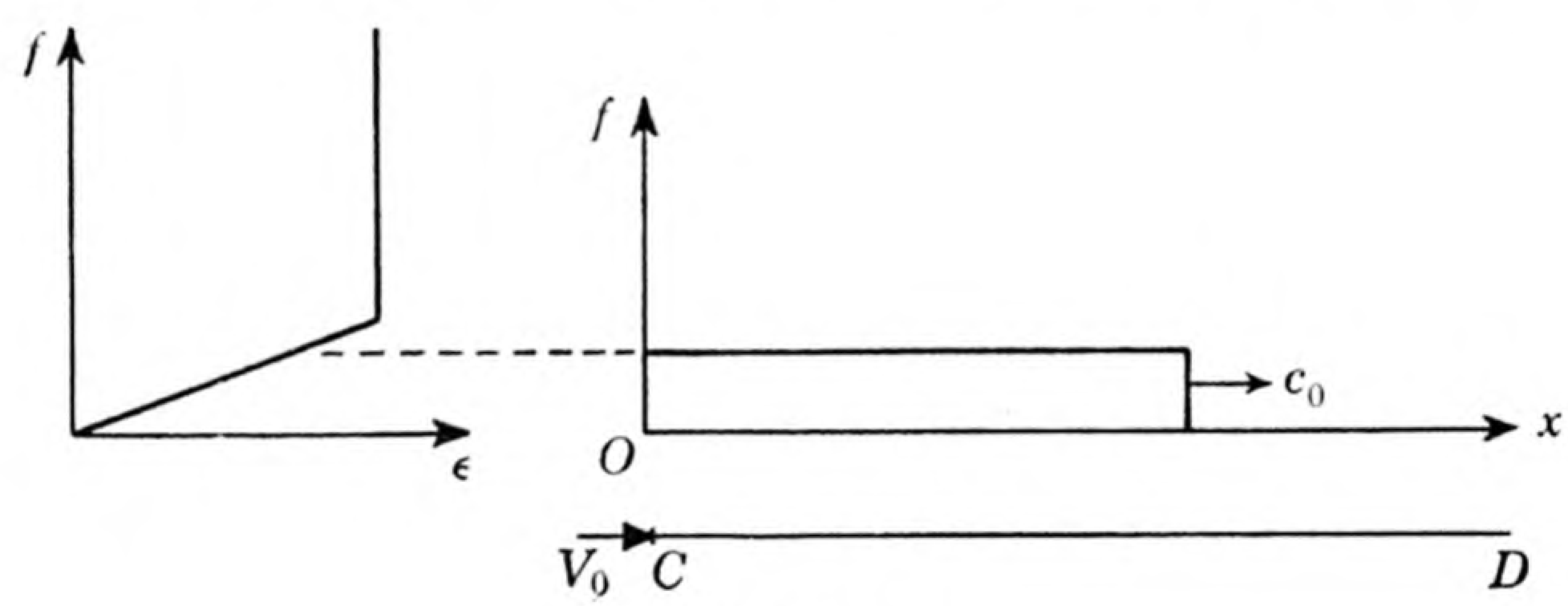


FIG. 4

it propagates. The relation for a spring with closure depicted in Fig. 3 is concave upward, and the resulting modification of the theory, suggested by White and Griffis [6] and by Lee [7] must be adopted. In this case a wave of increasing stress may have a continually steepening wavefront, and a shock wave of force discontinuity may arise.

Let us consider some simple boundary-value problems which provide examples of the types of behavior which this theory predicts. Figure 4 illustrates the problem of a semi-infinite spring CD, $0 \leq x$, initially at rest and unstressed, and subjected to a constant end velocity V_0 at zero time. If V_0 is sufficiently small, as shown in Fig. 4, coil closure will not occur and the well-known linear elastic theory developed in the first section applies. A wave of constant material velocity V_0 and constant force and strain is propagated down the spring with velocity c_0. The continuity equation determines

$$V_0 = c_0\epsilon \tag{8}$$

and momentum considerations at the wavefront determine

$$f = mc_0V_0. \tag{9}$$

This solution will apply if the strain given by (8) is less than the closure strain $\epsilon_{\max}$.

For a larger value of prescribed end velocity, say V_1, (8) determines a strain greater than closure strain, and the solution shown in Fig. 5 applies. A wavefront of spring closure travels down the spring, producing conditions corresponding to the point B in the (f,ϵ)-plane. Behind this front the coils are in contact and moving with the prescribed end velocity V_1. The continuity equation and the known closure strain ϵ_{max} determine the wavefront velocity c according to the relation

$$V_1 = c\epsilon_{max}, \tag{10}$$

and momentum considerations at the wavefront determine the closure force by

$$f = mcV_1. \tag{11}$$

Thus the velocity c of the closure wave increases linearly with V_1 and the force linearly with $(V_1)^2$. It is interesting to observe that the elastic part of the force-strain relation does not influence this solution as long as closure

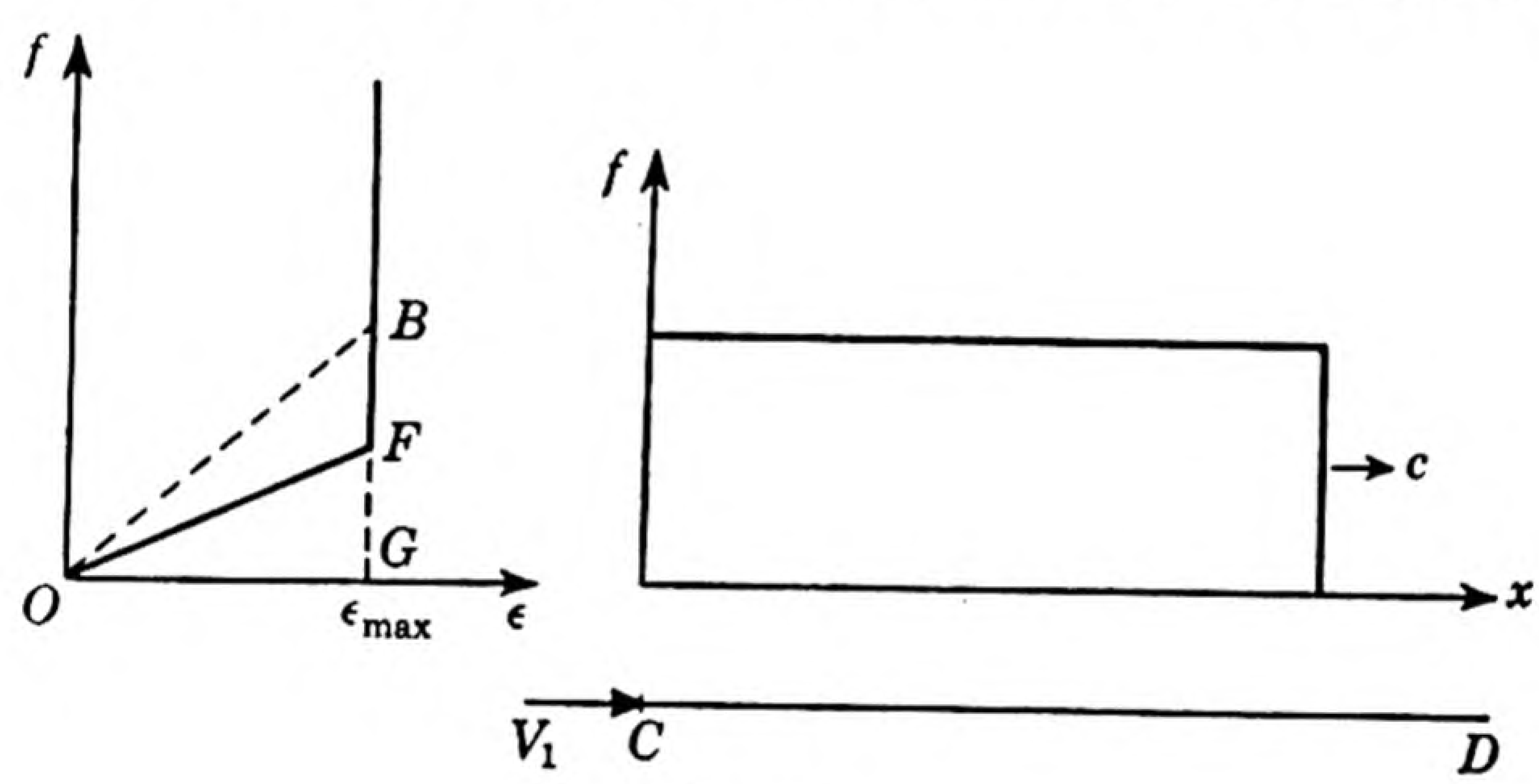


FIG. 5

occurs, but it does influence the magnitude of the impact at closure. The force required to overcome the elastic restraint in order to initiate closure transmits part of the impulse at the wavefront, and only the additional closure force is transmitted through the coil contacts.

In the case of impact with velocity V_0, which does not cause closure, energy is conserved, and the work done at the point of impact is divided equally between kinetic energy and elastic-strain energy of the spring. In the case of impact causing coil closure depicted in Fig. 5, half the work done at the point of impact appears as kinetic energy; the rest represents work done on the spring and is equivalent, per unit length of spring, to the area OBG in Fig. 5. Only the portion OFG occurs as recoverable elastic energy, the part OBF being lost owing to the inelastic coil-on-coil collision.

The present application of the theory of longitudinal waves in a rod with a concave upward stress-strain relation is more satisfactory physically than the previously available application of plastic waves in compression. The nominal-stress nominal-strain relation for a metal in compression is of the form shown in Fig. 6a. There is a vertical asymptote at the strain unity, since unit strain implies reduction of the length of the rod to zero, which would corre-

spond to squeezing the rod down to a flat sheet. The associated increase in area demands that the nominal stress increases without limit, so that a vertical asymptote occurs, and the stress-strain relation to be used in this type of theory has a concave-upward region. A longitudinal compressive wave in such a material, analyzed on the basis of longitudinal motion only, leads to a wave of stress and strain discontinuity, represented, for example, by the jump AB shown in Fig. 6a. Thus a wave of discontinuous change in section as shown in Fig. 6b is predicted. Such a wave would imply infinite lateral acceleration due to the lateral expansion associated with the longitudinal compres-

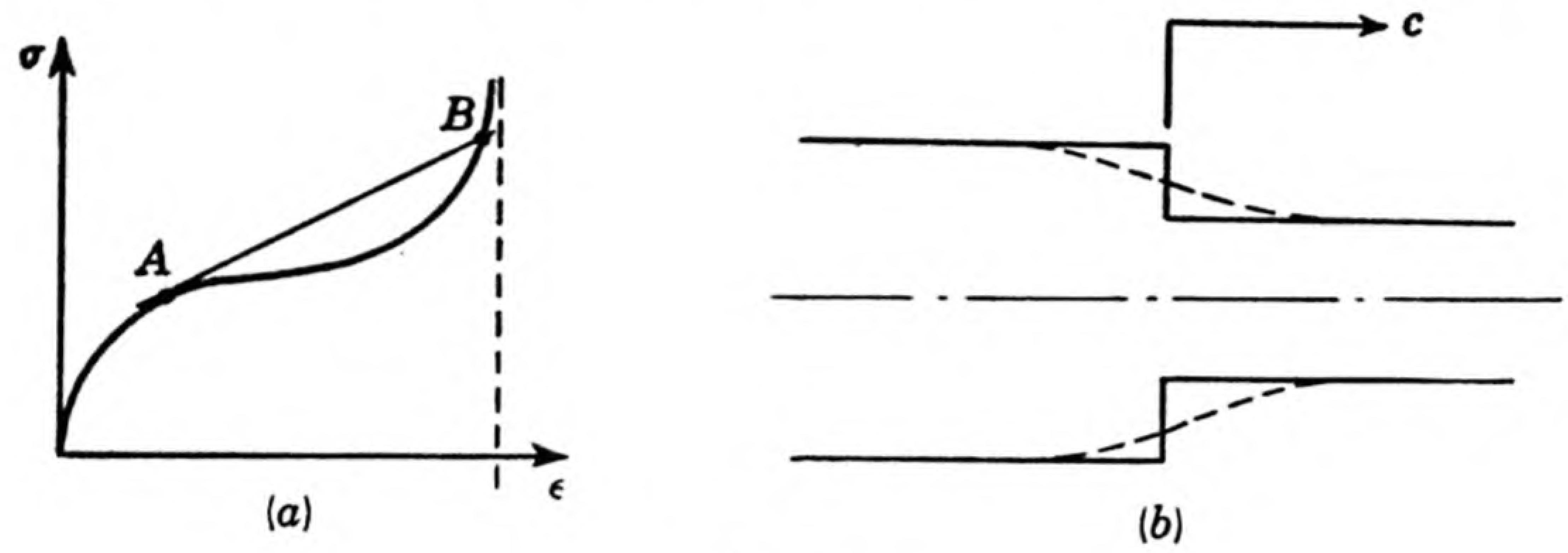


FIG. 6

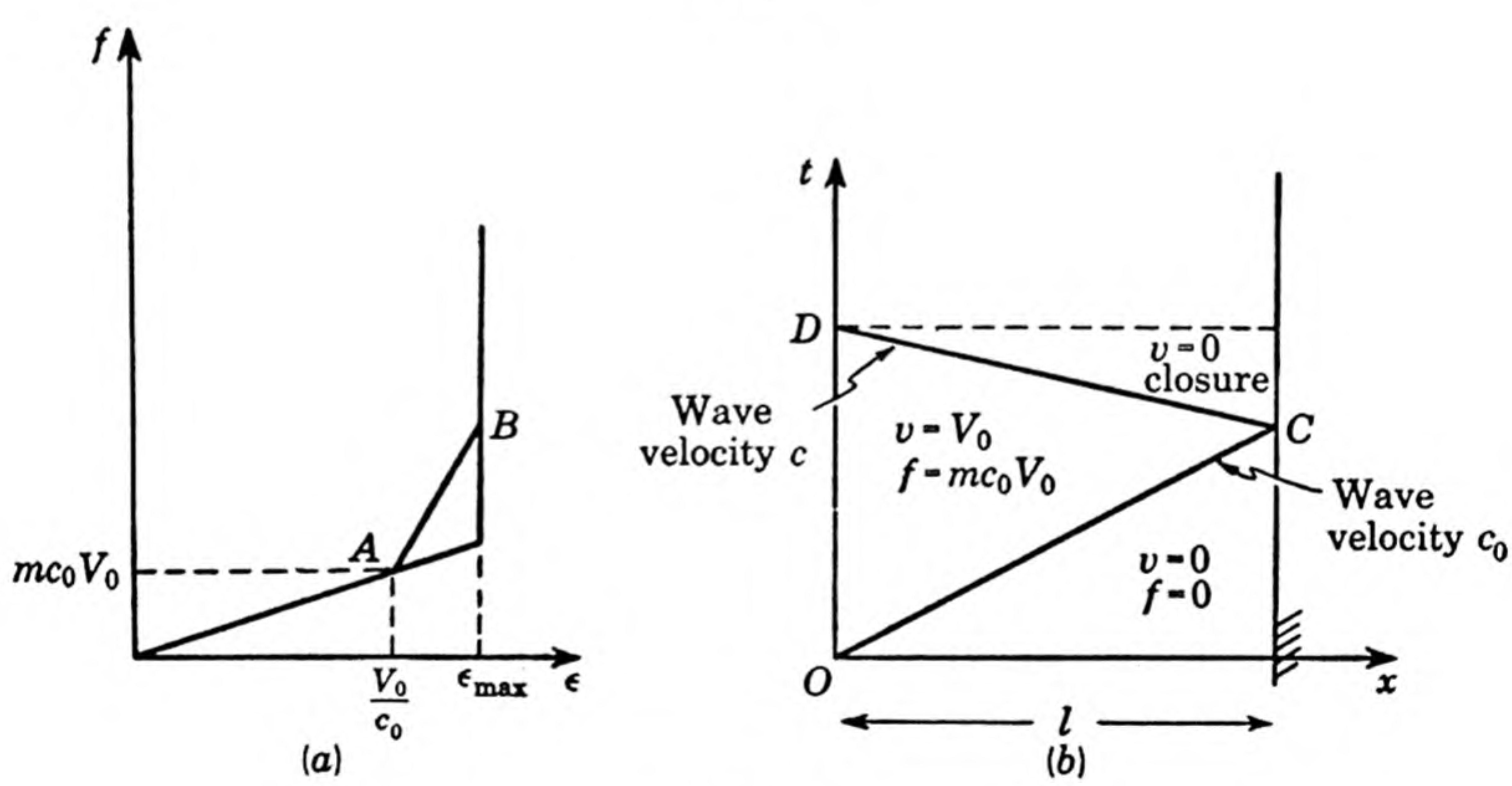


FIG. 7

sion, and lateral-inertia effects would have an important influence in smoothing out such a wavefront, as indicated by the broken lines in Fig. 6b, spreading it over several rod diameters. In the present application to spring-closure waves no such lateral-inertia effects arise, and this types of analysis can be expected to provide a closer approximation to the physical situation.

The application to more complex boundary-value problems can be readily treated by considering the wavefront configurations in the (x,t)-plane. Consider, for example, the problem shown in Fig. 7 of a constant end velocity V_0 applied to a spring of length l, the other end of which is fixed. Suppose V_0 does not cause closure because of the initial wave emanating from the impact end. Let A in Fig. 7a represent the stress and strain values behind this initial wavefront OC in Fig. 7b. Reflection at the fixed end at C based on a

linear elastic relation would produce a reflected wave of twice the stress amplitude of the initial wave, and the situation shown in Fig. 7 represents a case when such an elastic reflected wave would determine a strain greater than the closure strain. The reflected wave will then be a closure wave CD represented by the jump AB in Fig. 7a. The conditions to be satisfied are that closure occurs behind the wavefront CD so that the particle velocity v there is zero to satisfy the fixed-end condition, and the continuity condition across CD becomes

$$V_0 = c(\epsilon_{\max} - \epsilon_A) = c\left(\epsilon_{\max} - \frac{V_0}{c_0}\right). \tag{12}$$

Momentum considerations across the closure wave CD give

$$f_B - f_A = mcV_0 = f_B - mc_0V_0. \tag{13}$$

(12) determines the wave velocity c, which gives the gradient of CD in Fig. 7b, and (13) then determines the closure force f_B. When the closure wave reaches the impact end at D, the spring is closed completely and compatibility with the fixed-end condition demands that the velocity at the impact end should fall instantaneously to zero.

More complex boundary-value problems can be handled most easily by using the characteristic relations in the (x,t)-plane,

$$f \pm mc_0v = \text{const.} \qquad \text{on} \qquad x \mp c_0t = \text{const.} \tag{14}$$

when closure does not occur. Wavefronts of closure satisfy the continuity and momentum conditions,

$$\begin{aligned} \Delta v &= c\,\Delta\epsilon, \\ \Delta f &= mc\,\Delta v, \end{aligned} \tag{15}$$

where Δ represents the discontinuity of the variable due to the passage of the wave and c is the wave velocity. Combination of these relations with the closure condition $\epsilon = \epsilon_{\max}$ will determine the solution for prescribed boundary conditions.

3. Closure with perfectly elastic coil on coil rebound. We shall now consider the case in which perfectly elastic rebound occurs if adjacent coils collide with a nonzero relative velocity of approach. The time of contact, which will be given by the Hertz theory of impact [8], will be extremely short compared with the time of wave propagation along the spring. Thus such impacts will occur as pairs of curves along one of which a moving-impact point force is acting and along the other an equal and opposite force. These curves will be separated by the pitch distance p in the x-direction, since impact collision occurs in adjacent coils initially at a distance p apart. Since we are not considering the detailed propagation around a single coil and we are replacing the actual coil-closure condition (4) based on the relative displacement of sections initially a pitch length apart by the maximum-strain condition (5), such pairs of forces can be replaced by a force doublet containing equal and opposite

forces moving down the spring in close proximity to each other. Apart from these force doublets, the spring motion will be governed by the linear wave equation as discussed in the first section, so that the analysis of a spring with bouncing coil closure is replaced by the analysis of a linear wave system with an appropriate distribution of external forces. In the (x,t)-plane, the solution will consist of the trajectories of force doublets, between which the linear wave equation will be satisfied. In order to visualize how such solutions are developed, it is convenient to study an associated system with lumped masses.

Let us consider a row of equal rigid balls of diameter d sliding freely in a straight groove and initially separated uniformly from each other by the distance δ. Such a system can be considered as a simple analogue of a spring. The balls can approach each other through the distance δ, at which instant contact occurs and a force of any magnitude can be transmitted beween the balls. This model is equivalent to a spring with no elastic response before closure occurs. Let us consider a semi-infinite row of such balls and give the end one a constant velocity V leading to closure. For simplicity we first consider the case of inelastic impact discussed in the previous section. After a time δ/V, collision between the first and second balls will occur, and thereafter the second ball will continue to move with the velocity V, and the two balls will remain in contact. After a time $n(\delta/V)$, where n is an integer, the $(n + 1)$st ball will begin to move in contact with the first n balls. The wave-propagation velocity along the row of undisturbed balls is given by

$$c = (d + \delta)\frac{V}{\delta} \tag{16}$$

since the wavefront progresses a distance $d + \delta$ in each time interval δ/V. This corresponds to the velocity based on the Lagrange coordinate x for the spring. In fact the closure strain is

$$\epsilon_{\max} = \frac{\delta}{d + \delta} \tag{17}$$

and (16) is a repetition of (10). In the present case of lumped masses, momentum considerations will determine a series of impulses to maintain the constant velocity of the first ball in place of the constant force for the case of continuously distributed mass. The motion of the system of balls is thus seen to be closely related to the analysis given in the previous section for inelastic coil collision.

Let us now consider the case of the balls for perfectly elastic impact. If the first ball of the series is prescribed to move with the constant velocity V, after a time δ/V it strikes the second ball, which bounces forward with velocity $2V$ in order to maintain the same magnitude of the relative velocity on approach and separation. When this strikes the third ball after a further time interval $\delta/2V$, the velocities of the two balls are exchanged, according to the behavior in collision of perfectly elastic balls of equal mass. Thus the second ball is brought to rest, and the third ball is propagated forward with velocity $2V$.

This procedure is repeated, each ball being subjected to an impulse which projects it forward with velocity $2V$ and then a reverse impulse on striking the ball ahead which brings it to rest again. The first ball continues its forward motion with velocity V and again strikes the second ball, initiating a second double-impulse sequence which travels down the row of balls. The process is repeated as illustrated in Fig. 8, in which the abscissa is chosen in

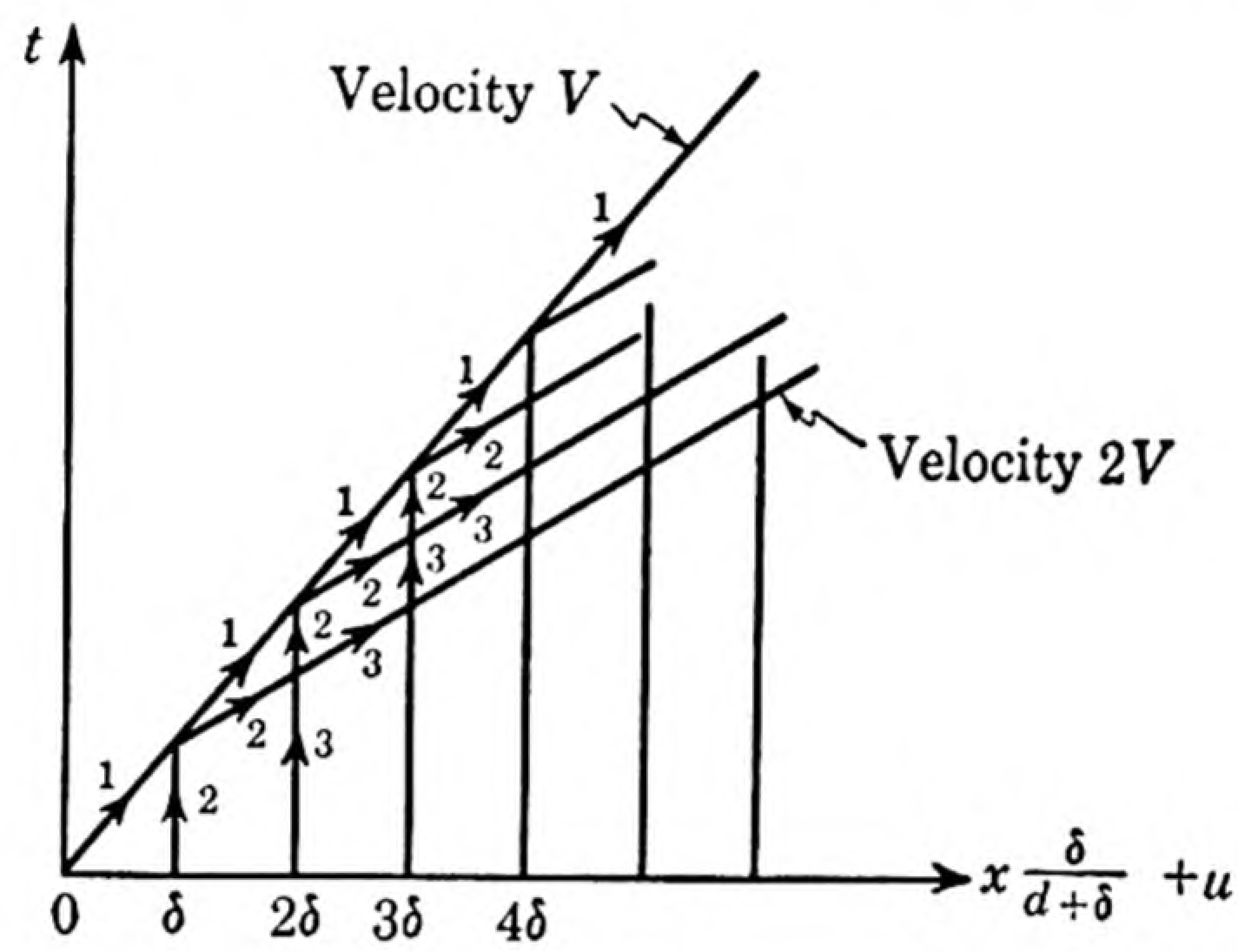


FIG. 8

terms of the initial position x and the displacement u to represent the balls as moving points. Since the velocity of the ball at the head of the wavefront is $2V$, the propagation velocity down the undisturbed row of balls is given by

$$c = 2(d + \delta) \frac{V}{\delta}. \tag{18}$$

Behind this wavefront the balls are alternately at rest or in motion with the velocity $2V$, apart from the first ball, which moves with the constant velocity V. Since the balls behind the wavefront are at rest and in motion during equal intervals of time, their average velocity is V and they move forward on the average in step with the first ball. Figure 8 shows clearly that the wave is propagated by a series of impulse doublets which move along the undisturbed line of balls with the velocity c of (18) and produce successively on each ball first a forward and then a backward impulse. This sequence provides a model for the consideration of the corresponding problem of spring surges. For the spring problem with a continuous distribution of mass, the moving impulse pair will be replaced by a moving force doublet, and we must investigate the response of the spring to such a force system.

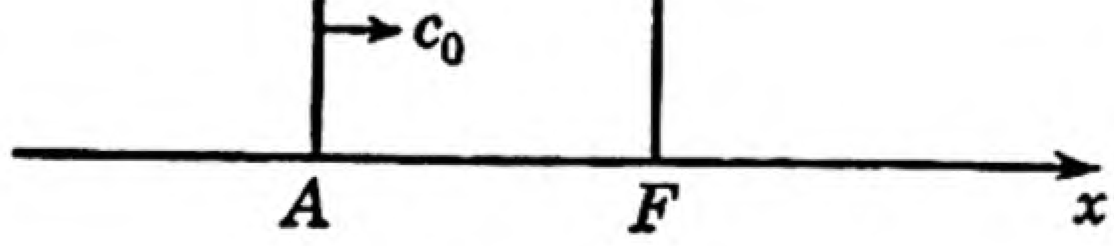


FIG. 9

Figure 9 illustrates the stress or velocity distribution associated with a force doublet moving through an initially undisturbed spring. Since the doublet is produced by coil closure and rebound, it will move with a velocity

c which is faster than the elastic wave velocity c_0. Thus no disturbance can be propagated ahead of the forward force pulse EF in Fig. 9, and a wavefront of stress and velocity discontinuity arises. Since the forces are produced by closure impact, the strain in the region DE of coil contact must be the closure strain ϵ_{max}. Since we can replace the effect of coil closure by that of moving forces on a system which satisfies the linear wave equation, superposition can be used and the wavefront EF would simply be repeated with a negative sign to represent the motion due to the negative component of the force doublet at D. However, when we consider the associated force-strain relation for a spring with coil closure illustrated in Fig. 3, the concave-upward property suggests shock-wave conditions on loading and a spreading unloading wave having a tail propagated with the smaller velocity c_0. The situation is therefore examined in detail below, and it is shown that the region BC does in fact correspond to zero stress and velocity, so that nonzero stress and velocity occur only within the force doublet. Let f and v be the spring force and velocity along DE, and f_1 and v_1 along BC. The momentum relation across the wavefront EF is similar to (15) but contains an additional term due to the external force T. It becomes

(19) $$f + T = mcv.$$

Continuity across EF is equivalent to the first of (15),

(20) $$v = c\epsilon_{max}.$$

But f is the elastic closure force $mc_0^2\epsilon_{max}$, since additional contact forces are taken care of by the force doublet. Thus, for a prescribed value of v, T and c are determined.

For the first part of the unloading wave DC, the momentum relation is

(21) $$(f - f_1) + T = mc(v - v_1),$$

the continuity equation

(22) $$(v - v_1) = c(\epsilon_{max} - \epsilon_1),$$

and f_1 and ϵ_1 must be related by the force-strain law for the spring,

(23) $$f_1 = mc_0^2\epsilon_1.$$

Subtracting (21) from (19) gives

(24) $$f_1 = mcv_1,$$

and (22) from (20)

(25) $$v_1 = c\epsilon_1.$$

(24) and (25) combine to determine

(26) $$f_1 = mc^2\epsilon_1.$$

Thus, for $c > c_0$, (23) and (26) are incompatible unless $f_1 = \epsilon_1 = 0$. Thus a force doublet produces a pulse of force, strain, and velocity in the spring which moves along with it, and B and C in Fig. 9 lie on the x-axis. Thus the rebound conditions on coil closure are such that the force and velocity conditions ahead of the closure are regained when it has passed. We note that as for the case of inelastic coil collision, the elastic force-strain relation does not influence the speed of propagation of the closure waves, but it does affect the magnitude of the impact force T. We shall now apply these conditions to simple boundary-value problems for springs with bouncing coils.

We shall first consider a semi-infinite spring, $x \geq 0$, initially undisturbed, subjected to constant-velocity impact at the end $x = 0$. The impact velocity

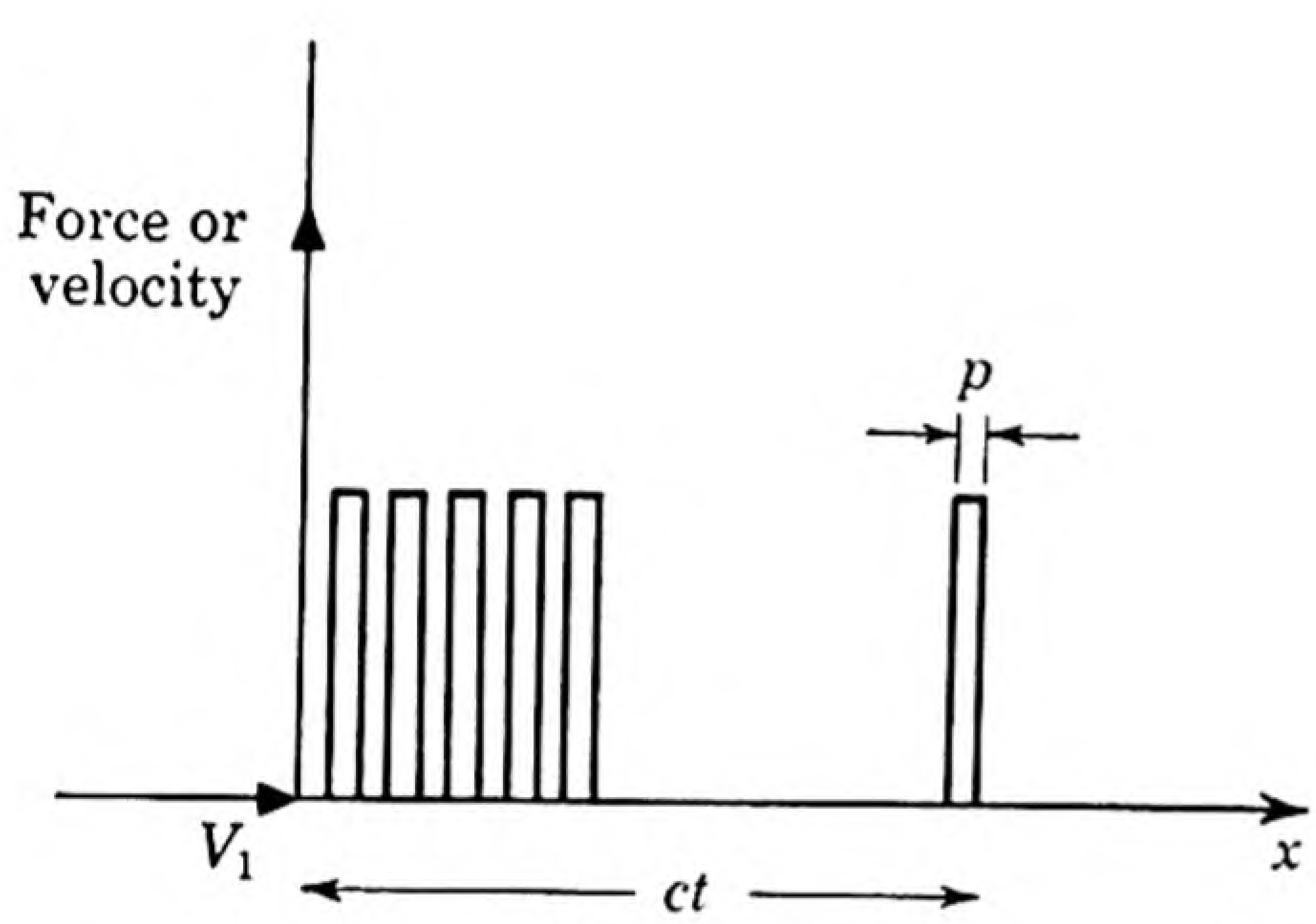


FIG. 10

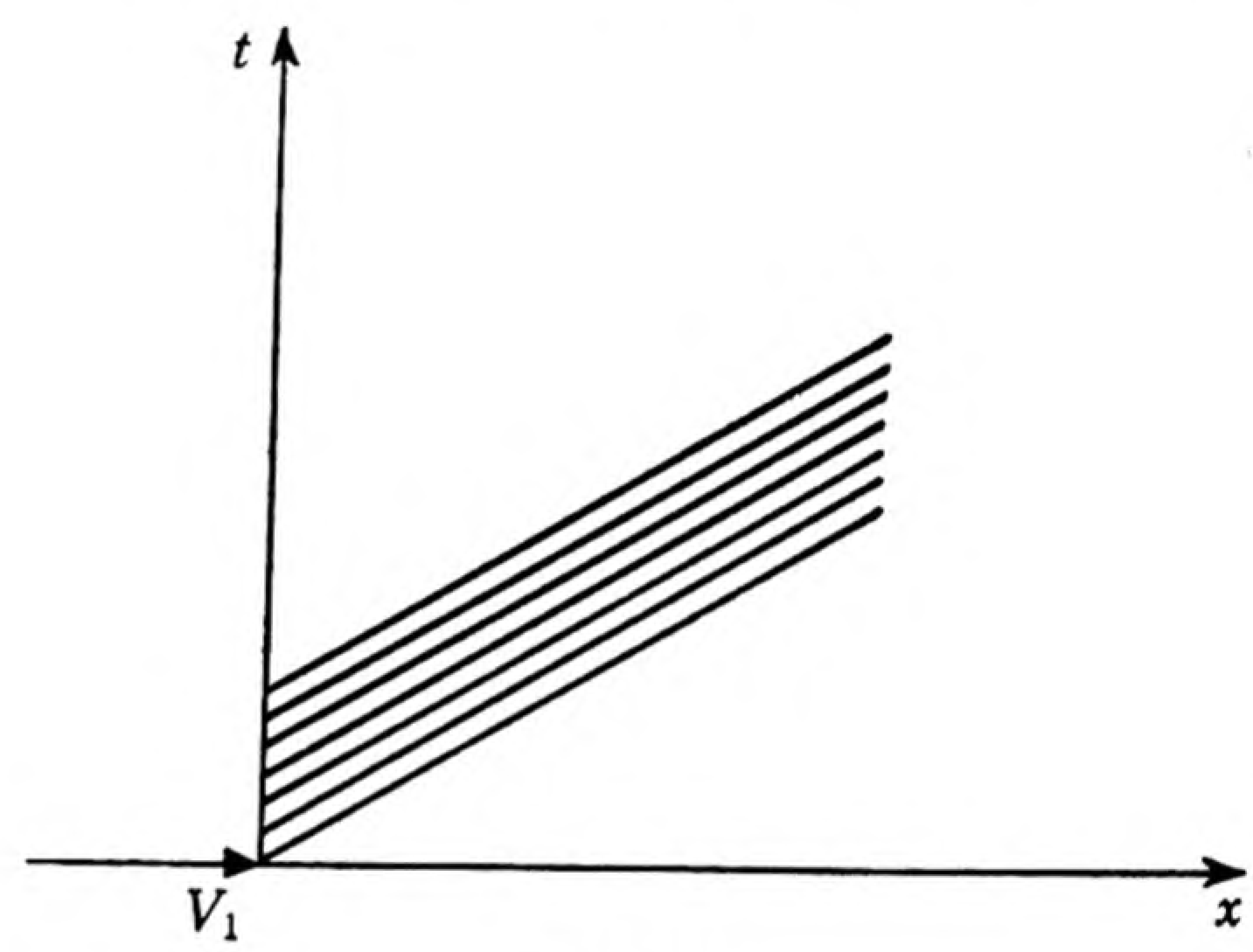


FIG. 11

V_1 is sufficiently large to cause coil closure, so that substitution in (8) would indicate an elastic strain greater than ϵ_{max}. When the first coil of the spring, which is constrained to move with the impact velocity V_1, strikes the next coil, the latter rebounds forward with velocity $2V_1$ because of the perfectly elastic assumption of coil-on-coil collision. This material velocity $2V_1$ determines the magnitude of the force doublet which represents the advancing closure front. The wavefront velocity c is given by (20) with $v = 2V_1$ and the force-doublet magnitude by (19). This first doublet is followed by a sequence of identical force doublets generated by the constant-velocity boundary condition since the spring is left unstrained and at rest after each doublet has passed. The force and velocity distribution are therefore as depicted in Fig. 10, and the (x,t)-plane configuration is as shown in Fig. 11. Since the boundary velocity is V_1 and the wavefront speed $2V_1/\epsilon_{max}$, the average strain behind the wavefront is $\epsilon_{max}/2$. Thus the doublets must be spaced so that equal intervals of x are alternately strained to closure and unstrained. According to our theory in which the actual closure condition (4) is replaced by the maximum-strain condition (5), the strain pulses are infinitesimal in extent, depending on the Hertz time of impact. However, in the actual situation, the coil-collision forces are separated by a pitch length p in x, so that after an element of the spring has been set in motion by coil collision it will not be brought to rest

again until the wave has traveled a pitch length along the spring. This determines the scale of the velocity pulses illustrated in Fig. 10.

Let us now consider a finite spring, $0 \leq x \leq l$, subjected to constant-velocity impact at $x = 0$, and held fixed at $x = l$. Suppose that the impact velocity V is such that closure first occurs on reflection at the fixed end. Figure 12 shows the (x,t)-plane configuration. AO is the elastic wavefront, and the spring force and strain behind this front are represented by the point D in the force-strain diagram (Fig. 13). At A a reflected wave must be developed which will maintain the end of the spring at rest. The relative velocity of collision between the coil adjacent to A which is set in motion by the advancing elastic wave and the coil held fixed at A is V, so that elastic rebound determines a rebound velocity $-V$ along the reflected wavefront AC. This magnitude

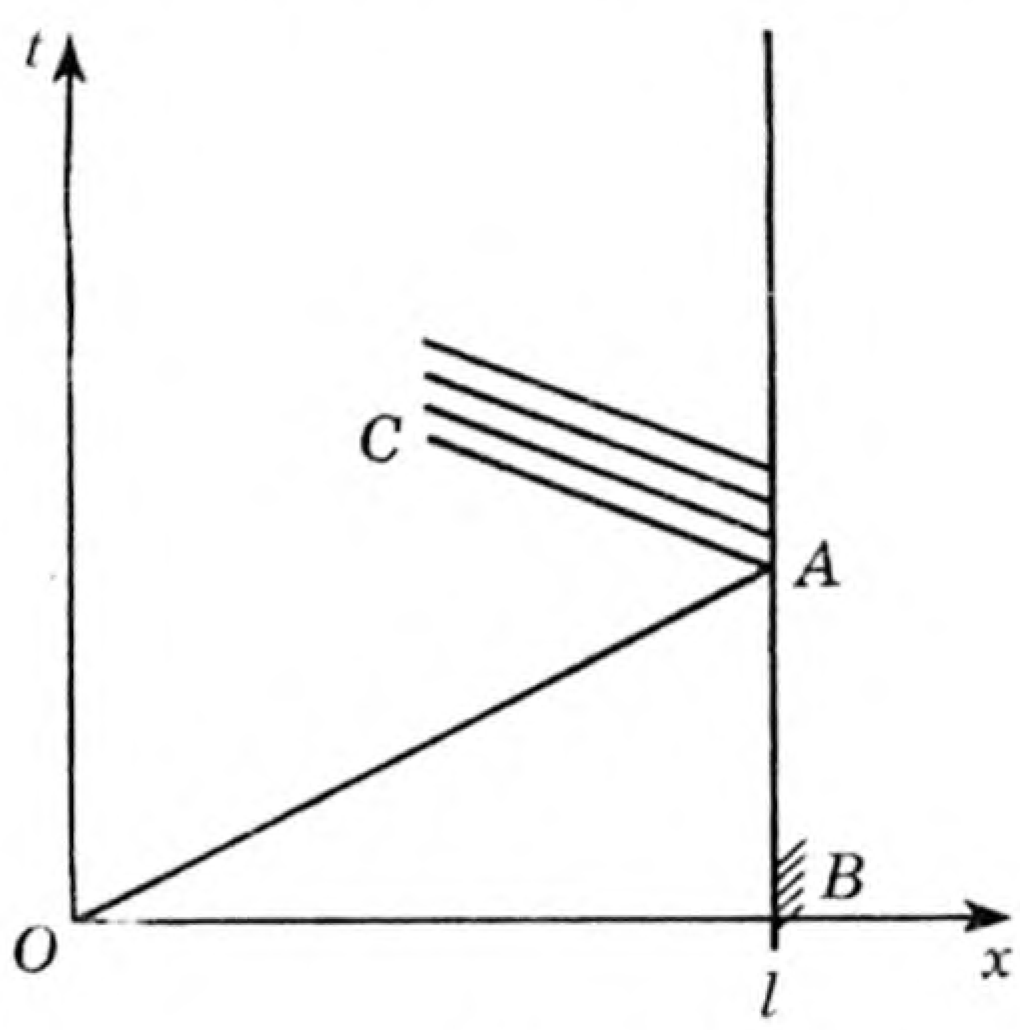


FIG. 12

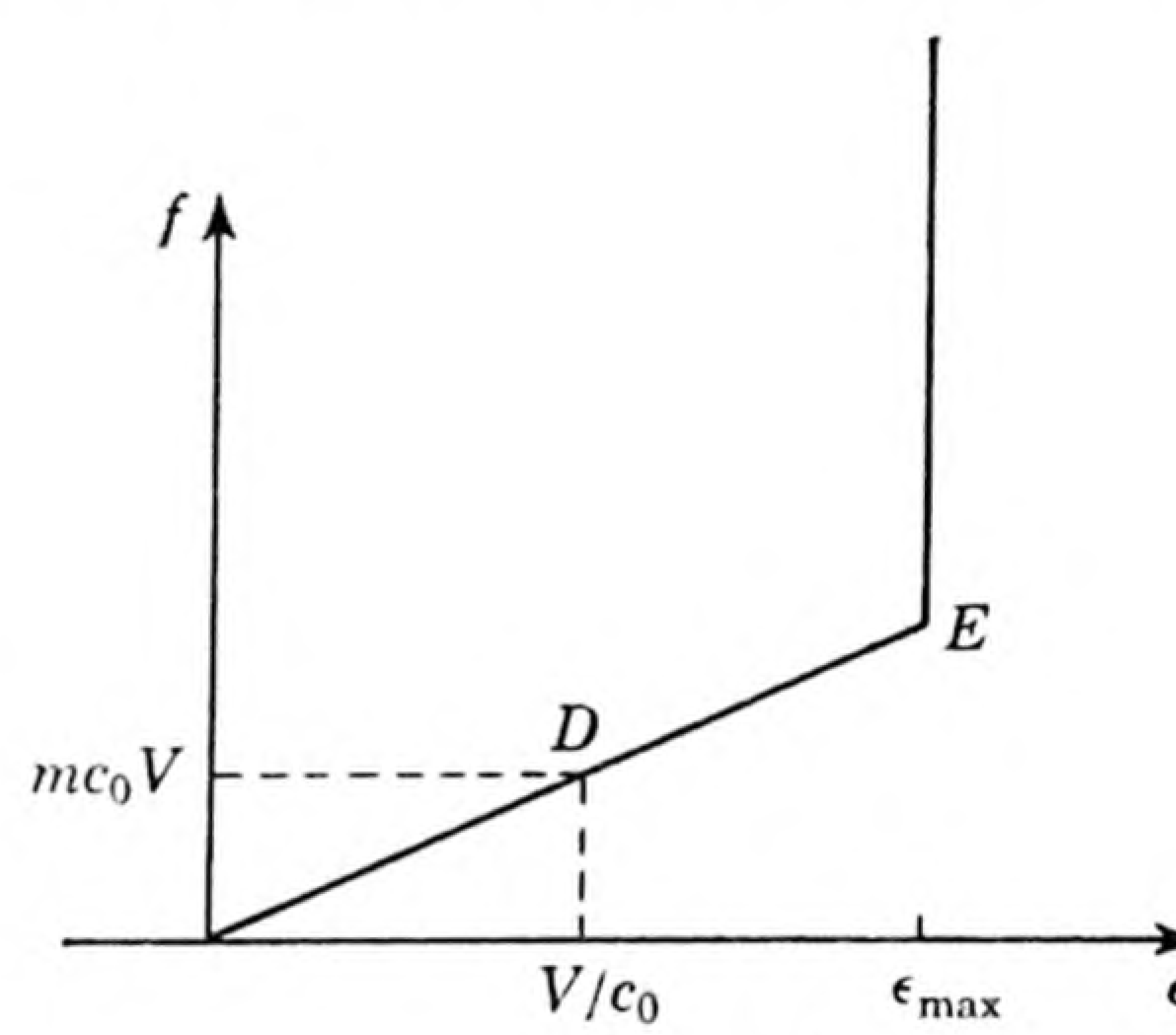


FIG. 13

will determine the force doublet due to closure and the speed of propagation of AC. With a relative velocity at the wavefront of $2V$, continuity gives the wave speed c according to

$$2V = c\left(\epsilon_{\max} - \frac{V}{c_0}\right). \tag{27}$$

Momentum considerations give the impact force according to

$$T + f_E - f_D = mc2V, \tag{28}$$

where f_E is the elastic closure force corresponding to E in Fig. 13 and f_D the spring force in the region OAC of Fig. 12. The wavefront AC will be associated with a doublet, behind which the conditions in OAC will be regained. A sequence of such doublets will be produced by reflection at the fixed end, and as in the previous case the average change in strain behind AC is half that at the doublet, so that regions of closure occupy half the spring length traversed, the intermediate intervals maintaining the conditions in OAC. Further reflection will occur at the impact end, and the detail analysis of this will require a study of the interaction between the force doublets produced by reflection at each end, which is not treated in the present paper. More complex boundary-value problems can be treated by a development of this type of analysis.

4. Conclusions. The theory developed above forms a basis for the analysis of spring surges with coil closure for both inelastic and elastic collision. The two cases dealt with, inelastic and perfectly elastic collision, determine contrasting solutions. The rebound momentum due to coil bounce causes much higher wave-propagation velocities for the same impact velocity, the ratio being 2:1 for a simple closure wave in an initially undisturbed spring. While it is felt that the present type of analysis will be satisfactory, the assumptions involved in the consideration of simple longitudinal motion, which are discussed in the introduction, cannot at present be assessed theoretically. Experimental check of the theory would therefore be very valuable. This is particularly so in the case of elastic rebound, since the motion consists of pulses, each a coil pitch in wavelength, and it is on this scale that the detail of propagation round the helix may influence the motion. However, the fact that the simple elastic theory agrees with experiment [4] for a wavefront of velocity discontinuity, for which experimentally a wavefront exhibiting little dispersion was obtained, lends support to the suggestion that the type of theory developed in this paper will form a satisfactory method of analysis of spring surges with coil closure.

Bibliography

1. K. J. Dejuhasz, *Graphical analysis of surges in mechanical springs*, J. Franklin Inst. vol. 226 (1938) pp. 505–526.

2. H. G. Woodhouse Taylor, *Wave action in gun run-up springs*, Proc. Inst. Mech. Engrs. London vol. 145 (1941) pp. 150–159.

3. A. E. H. Love, *Propagation of waves of elastic displacement along a helical wire*, Trans. Cambridge Philos. Soc. (1899).

4. W. Weibull, *De Dynamiska Egenskaperna hos Spiralfjadrar*, Svenska Bokhandelscentralen A-B Stockholm (1927).

5. T. von Kármán, H. F. Bohnenblust, and D. H. Hyers, *The propagation of plastic waves in tension specimens of finite length*, NDRC Rep. A-103 (O.S.R.D. No. 946) (1942).

6. M. P. White and L. Griffis, *Wave propagation in a uniform bar whose stress-strain curve is concave upward*, NDRC Rep. A-152 (O.S.R.D. No. 1302) (1943).

——— and ———, *The propagation of plasticity in uniaxial compression*, J. Appl. Mech., Amer. Soc. Mech. Engrs. Trans. vol. 70 (1948) pp. 256–260.

7. E. H. Lee, *Plastic waves in compression*, British Official Rep. A.P.P. Coord. Sub-Committee No. 57 (1943).

8. H. Hertz, J. Reine Angew. Math. vol. 92 (1881) p. 156.

Brown University,
Providence, R.I.

ON THE WAVE EQUATION AND THE EQUATION OF EULER-POISSON[1]

BY

ALEXANDER WEINSTEIN

1. Introduction. A symposium on wave propagation is a suitable opportunity to present some results on a remarkable class of hyperbolic partial differential equations which includes the classical wave equation as a special case. By considering this class of equations we shall obtain a new and perhaps the simplest solution of Cauchy's problem for the wave equation. But it should be emphasized that this is not the sole purpose of presenting this paper. In fact, some of the equations considered here, such as the Tricomi equation, are known to have important applications in other fields such as transonic flow of compressible fluids. Moreover, our hyperbolic equations present a remarkable counterpart to a class of elliptic equations in generalized axially symmetric potential theory which also give rise to a number of applications in hydrodynamics and the theory of elasticity.

2. The Euler-Poisson-Darboux equation. Let x be an abbreviated notation for a system of m variables $x_1, x_2, \cdots, x_m$, $m \geq 1$. In the following we write $u(x,t)$ for $u(x_1,x_2, \cdots ,x_m,t)$. Let k be any real number, $-\infty < k < \infty$. We consider the hyperbolic differential equation

$$\Delta u^k = u^k_{tt} + kt^{-1}u^k_t, \tag{1}$$

where the superscript k indicates the dependence of u on the parameter k. We write sometimes $u^{(k)}$ in place of u^k. The symbol Δ is an abbreviation for $\sum_{i=1}^{m} (\partial^2/\partial x_i^2)$. We intend to give here a solution of Cauchy's problem for (1) with the initial conditions

$$u^k(x,0) = f(x_1,x_2, \cdots ,x_m), \qquad u^k_t(x,0) = 0, \tag{2}$$

for a given sufficiently regular function f. The special case $k = 0$ yields a wave equation, $\Delta u^0 = u^0_{tt}$. Let us observe that, except for $k = 0$, equation (1) has variable coefficients. Moreover, we are considering here a singular initial-value problem since the coefficient kt^{-1} is infinite on the hyperplane $t = 0$. This implies that the usual existence and uniqueness theorems for Cauchy's problems are not applicable to our case.

For $m = 1$, equation (1) appears already in Euler's work. Actually, Euler considers an even more general equation containing, besides k, a second parameter. Euler's case was discussed in a famous paper by Poisson. A brilliant exposition of Poisson's theory, always for $m = 1$, was given by Darboux [1]. In our notation Darboux considered only the range $0 < k < 2$, under the

[1] This paper was sponsored by the Office of Naval Research.

mistaken impression that no essential new features would appear outside this range. Incidentally, the case $k = \frac{1}{3}$ corresponds to Tricomi's equation. Besides the investigation of Euler's equation Poisson has given in another memoir his famous solution of Cauchy's problem for the wave equation in ordinary three-dimensional space by using (1) for $m = 3$ and $k = 2$. It seems that the two cases $m = 1$, $0 < k < 2$, and $m = 3$, $k = 2$, appeared unrelated in the classical work of Poisson.

The cases $m = 2, 3, \cdots, k = m - 1$, and $m = 1$, $k = 1, 2, 3, \cdots$, were considered by Asgeirsson, who calls (1) the Darboux equation, a terminology which was retained in Courant-Hilbert [2, p. 381], while Darboux, who in fact considers only the case $m = 1$, calls (1) the Euler-Poisson equation. The cases $m = 1$, $k = -1, -2, -3, \cdots$, were considered by M. H. Martin [3] and J. B. Diaz and M. H. Martin [4], in a remarkable extension of Riemann's method. M. B. Kapilevic [5] has recently considered (1) for $0 < k < 1$ and $m = 1$ and $m = 2$. Actually, the equation considered by Kapilevic, although seemingly more general, can be obtained from (1) by Hadamard's method of descent [7]. The solution of Cauchy's problem given by Kapilevic is very complicated and involves finite parts of integrals. In our exposition we shall consider Cauchy's problem for all positive integers m and all real values of k and shall call (1) the equation of Euler-Poisson-Darboux (abbreviated EPD). The general method presented here is an improvement on that sketched in a preliminary note by the present author [6]. Further investigations were made by Diaz and Weinberger [9] and Blum.

3. Fundamental recursion formulas. We have the following two fundamental recursion formulas:

$$u_t^k(x,t) = tu^{k+2}(x,t), \tag{3}$$

$$u^k(x,t) = t^{1-k}u^{2-k}(x,t). \tag{4}$$

These recursion formulas allow us to obtain from a solution u^k of equation (1) solutions of the same equation with the parameter $k + 2$ and $2 - k$, respectively. Both formulas are mentioned in a somewhat different form in the book of Darboux, but only for the case $m = 1$. As similar formulas occur in other parts of the theory of differential equations and as the proof is usually omitted, it is worth while to point out that such formulas are simple consequences of the following identities for functions v and w of a single variable t:

Put $w = t^{k-1}v$. Then

$$w_{tt} + \frac{2-k}{t} w_t = t^{k-1}(v_{tt} + kt^{-1}v_t). \tag{5}$$

On the other hand, if we put $tw = v_t$, we obtain

$$w_{tt} + \frac{k+2}{t} w_t = \frac{1}{t}\frac{\partial}{\partial t}\left(v_{tt} + \frac{k}{t} v_t\right). \tag{6}$$

In the following paragraphs we shall solve (1), (2) explicitly for $k \geq m - 1$. Then by a process which is essentially a repeated application of (3) and (4)

we shall obtain a formula for the solution of Cauchy's problem for the case $k < m - 1$.

4. Cauchy problem for the case $k = m - 1$. It is well known [2] that the Cauchy problem for the EPD equation

$$\Delta u^{m-1} = u_{tt}^{m-1} + \frac{m-1}{t} u_t^{m-1} \tag{7}$$

with the initial conditions

$$u^{m-1}(x,0) = f(x_1,x_2, \cdots, x_m), \qquad u_t^{m-1}(x,0) = 0 \tag{8}$$

admits a unique solution given by a generalization of a celebrated formula of Poisson.

Let

$$\omega_m = \frac{2(\sqrt{\pi})^m}{\Gamma(m/2)} \tag{9}$$

denote the surface of the unit sphere in the m-dimensional x-space. Then the solution of the Cauchy problem (1) and (2) is given by the formula

$$\begin{aligned} u^{m-1}(x,t) &= \frac{1}{\omega_m} \int \cdots \int f(x_1 + \alpha_1 t, x_2 + \alpha_2 t, \cdots, x_m + \alpha_m t)\, d\omega_m \\ &= M(x,t;f), \end{aligned} \tag{10}$$

where M represents the mean value of f taken on a sphere of center x and of radius t. In (10) $\alpha_1, \alpha_2, \cdots, \alpha_m$ denote the components of a unit vector in x-space. The proof of our statement can be easily checked by direct computations which are given explicitly in the book of Courant-Hilbert [2, p. 411]. Sometimes we shall write in place of (10) in abbreviated notation

$$u^{m-1} = M(x,t;f).$$

5. The method of descent and the cases $k = m, m + 1, \cdots$. Let us consider always for a fixed m the case in which the parameter k is one of the integers $m, m + 1, \cdots$. Let us imagine that for such a given k we have an equation of the type (1), *viz.*,

$$\frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_m^2} + \frac{\partial^2 u}{\partial x_{m+1}^2} + \cdots + \frac{\partial^2 u}{\partial x_{k+1}^2} = \frac{\partial^2 u}{\partial t^2} + \frac{k}{t}\frac{\partial u}{\partial t} \tag{11}$$

with the conditions

$$u(x_1, \cdots, x_{k+1}, 0) = f_1(x_1, \cdots, x_{k+1}), \qquad u_t(x_1, \cdots, x_{k+1}, 0) = 0, \tag{12}$$

but in which we have, besides t, $k + 1$ independent variables $x_1, \cdots, x_{k+1}$ in place of $x_1, \cdots, x_m$. This new Cauchy problem is solved using (10) by the formula

$$\begin{aligned} &u(x_1, \cdots, x_{k+1}, t) \\ &\quad = \frac{1}{\omega_{k+1}} \int \cdots \int f_1(x_1 + \alpha_1 t, x_2 + \alpha_2 t, \cdots, x_{k+1} + \alpha_{k+1} t)\, d\omega_{k+1}. \end{aligned} \tag{13}$$

Let us put

$$f_1(x_1,x_2, \cdots ,x_m,0, \cdots ,0) = f(x_1,x_2, \cdots ,x_m), \tag{14}$$

where f is the function which appears in the initial conditions (2). Let us moreover put

$$f_1(x_1,x_2, \cdots ,x_m,x_{m+1}, \cdots ,x_{k+1}) = f_1(x_1,x_2, \cdots ,x_m,0, \cdots ,0). \tag{15}$$

In this way the u of (14) becomes a function only of $x_1, x_2, \cdots, x_m$ which satisfies equation (1) and the initial conditions (2) for the corresponding integral value of $k \geq m$.

As the Cauchy problem (1) and (2) admits for such values of k only one solution (see Sec. 11), we can write in place of (13)

$$\begin{aligned} u &= u^k(x_1,x_2, \cdots ,x_m,t) \\ &= \frac{1}{\omega_{k+1}} \int \cdots \int f(x_1 + \alpha_1 t, \cdots ,x_m + \alpha_m t)\, d\omega_{k+1}. \end{aligned} \tag{16}$$

The method described here for obtaining u^k for $k = m, m + 1, \cdots$, is the celebrated method of descent introduced by Hadamard [7]. It is of interest for the following to rewrite the formula (16) by expressing the k-tuple surface integral as an m-tuple integral. For the convenience of the reader we shall sketch this reduction. As x and t appear in the k-tuple integral of (16) only as parameters, it will suffice to consider the k-tuple integral

$$\frac{1}{\omega_{k+1}} \int \cdots \int f(\alpha_1,\alpha_2, \cdots ,\alpha_m)\, d\omega_{k+1}, \tag{17}$$

which obviously can be written by projecting on the equatorial hyperplane as

$$\frac{2}{\omega_{k+1}} \int \cdots \int \frac{f(\alpha_1, \cdots ,\alpha_m)\, d\alpha_1 \cdots d\alpha_m\, d\alpha_{m+1} \cdots d\alpha_k}{\sqrt{(1 - \alpha_1^2 - \cdots - \alpha_m^2) - \alpha_{m+1}^2 - \cdots \alpha_k^2}}. \tag{18}$$

Let $R^2 = 1 - \alpha_1^2 - \cdots - \alpha_m^2$. Then (18) can be rewritten as

$$\frac{2}{\omega_{k+1}} \underset{\sum_{i=1}^{m} \alpha_i^2 \leq 1}{\int \cdots \int} \frac{f(\alpha_1, \cdots ,\alpha_m)\, d\alpha_1 \cdots d\alpha_m}{R} \underset{\sum_{m+1}^{k} \alpha_i^2 \leq R^2}{\int \cdots \int} \frac{R\, d\alpha_{m+1} \cdots d\alpha_k}{\sqrt{R^2 - \alpha_{m+1}^2 - \cdots - \alpha_k^2}}. \tag{19}$$

The second integral in (19) is of the same type as (18) for $f = 1$. Therefore, it is half of the area of a sphere of radius R in a space of $k - m + 1$ dimensions,

so that its value is $\frac{1}{2}R^{k-m}\omega_{k-m+1}$. Putting this value in (19) we obtain

(20)

$$\frac{\omega_{k-m+1}}{\omega_{k+1}} \int \cdots \int_{\sum_{i=1}^{m} \alpha_i^2 \leq 1} f(\alpha_1, \cdots, \alpha_m)(1 - \alpha_1^2 - \cdots - \alpha_m^2)^{(k-m-1)/2} \, d\alpha_1 \cdots d\alpha_m.$$

With this result we obtain the following expression for (16):

$$(21) \quad u^k(x,t) = \frac{\omega_{k+1-m}}{\omega_{k+1}}$$

$$\int \cdots \int_{\alpha_1^2 + \cdots + \alpha_m^2 \leq 1} f(x_1 + \alpha_1 t, \cdots, x_m + \alpha_m t)(1 - \alpha_1^2 - \cdots \alpha_m^2)^{(k-1-m)/2} \, d\alpha_1 \cdots d\alpha_m.$$

6. The generalized method of descent. Formula (21) has been established for $k = m, m + 1, \cdots$. However, the integral on its right-hand side converges for all real values of k in the range $k > m - 1$. We shall verify now that the function $u^k(x,t)$ defined by the right-hand side of (21) satisfies the differential equation (1) and the initial conditions (2) for all values of k which are greater than $m - 1$. Of course, we have to substitute in (21) for ω_s the value $2(\sqrt{\pi})^s/\Gamma(s/2)$ as in (9) even if s is not an integer. This extension of the solution to nonintegral values of k can be called the generalized method of descent.

Let us put for a moment

$$(22) \qquad U(x,t) = \frac{\omega_{k+1}}{\omega_{k+1-m}} u^k(x,t).$$

Then we have to consider in place of (21) the following formula written in a self-explanatory abbreviated form:

$$(23) \quad U(x,t) = \int \cdots \int f(x_i + \alpha_i t)(1 - \rho^2)^{(k-1-m)/2} \, d\alpha_1 \cdots d\alpha_m,$$

$$(\rho^2 = \alpha_1^2 + \cdots + \alpha_m^2).$$

We have

$$(24) \qquad U_t = \sum_{i=1}^{m} \int \cdots \int f_i \alpha_i (1 - \rho^2)^{(k-1-m)/2} \, d\alpha_1 \cdots d\alpha_m,$$

where f_i denotes the derivative of f with respect to the argument $x_i + \alpha_i t$. Integrating by parts, we obtain

$$(25) \quad U_t = \sum_{i=1}^{m} \frac{t}{k - m + 1} \int \cdots \int f_{ii}(x_i + \alpha_i t)(1 - \rho^2)^{(k-m+1)/2} \, d\alpha_1 \cdots d\alpha_m$$

since each of the $(m-1)$-tuple integrals obtained in this process is obviously zero. Differentiating (25), we further obtain

$$(26)\quad U_{tt} = \sum_{i,j=1}^{m} \frac{t}{k-m+1} \int \cdots \int f_{iij}(x_i+\alpha_i t)\alpha_j(1-\rho^2)^{(k-m+1)/2}\,d\alpha_1 \cdots d\alpha_m + \frac{U_t}{t}.$$

Since

$$\frac{\partial}{\partial \alpha_j}[\alpha_j(1-\rho^2)^{(k-m+1)/2}] = (1-\rho^2)^{(k-m+1)/2} - (k-m+1)\alpha_j^2(1-\rho^2)^{(k-m-1)/2}$$

and since $j = 1, \cdots, m$, we obtain from (26) by an integration by parts

$$(27)\quad U_{tt} = \frac{U_t}{t} + \sum_{i=1}^{m} \frac{1}{k-m+1} \int \cdots \int f_{ii}(x_i+\alpha_i t) [-m(1-\rho^2)^{(k-m+1)/2} + (k-m+1)\alpha_j^2(1-\rho^2)^{(k-m-1)/2}]\,d\alpha_1 \cdots d\alpha_m.$$

The m boundary integrals which should appear in (27) are easily checked to be zero. Using (25) and (27), we have

$$(28)\quad U_{tt} + \frac{k}{t}U_t = \sum_{i=1}^{m} \int \cdots \int f_{ii}(x_i+\alpha_i t) (1-\rho^2)^{(k-m+1)/2}\,d\alpha_1 \cdots d\alpha_m = \Delta U.$$

Reverting back to u^k as introduced by (21), we see from (28) and from the definition of U that u^k indeed satisfies equations (1) and (2) for $k > m-1$.

Let us note that for (10) and (21) to be solutions it is sufficient to assume that f has continuous second derivatives.

7. The case $k < m - 1$. In this section, it will be shown that, for $k < m-1$ but $k \neq -1, -3, -5, \cdots$, the Cauchy problem can be solved by means of the results of Sec. 6 and repeated application of the fundamental formulas (3) and (4).

We begin by choosing a positive integer n such that $k + 2n \geq m - 1$. Using the methods of Sec. 6, we solve the Cauchy problem for the function $u^{k+2n}(x,t)$ satisfying the initial conditions

$$(29)\qquad u^{k+2n}(x,0) = \frac{f(x_1, \cdots, x_m)}{(k+1)(k+3)\cdots(k+2n-1)}, \quad u_t^{k+2n}(x,0) = 0.$$

By (4) we have

$$(30)\qquad t^{k+2n-1}u^{k+2n} = u^{2-k-2n},$$

and, applying (3) n times, we obtain from (30)

$$(31)\qquad \left(\frac{\partial}{t\,\partial t}\right)^n (t^{k+2n-1}u^{k+2n}) = u^{2-k}.$$

Finally, applying (4) again, we obtain from (31) the fundamental formula for the solution of Cauchy's problem,

$$u^k = t^{1-k}\left(\frac{\partial}{t\,\partial t}\right)^n (t^{k+2n-1}u^{k+2n}). \tag{32}$$

It is clear from the above construction, which starts from u^{k+2n}, that u^k is a solution of the EPD equation. Further, by (32),

$$u^k(x,t) = (k+1)(k+3)\cdots(k+2n-1)u^{k+2n}(x,t) + Ctu_t^{k+2n} + O(t^2), \tag{33}$$

where C is a constant. Therefore, by (29), u^k also satisfies the initial conditions (2).

Let us recall that for u^{k+2n} to be a solution it is sufficient that f have continuous second derivatives. In order to be able to carry out the construction of u^k, it is sufficient to require that f has n more derivatives. Thus, it is enough to assume that f has not less than $\frac{1}{2}(m - k + 3)$ continuous derivatives.

If $m - k$ is a positive odd integer, then n can be chosen so that

$$k + 2n = m - 1.$$

Therefore, u^{k+2n} in (32) is given by a Poisson mean-value formula of the type (10). In this case, and only in this case, we have a *Huygens' principle* as already noted [9].

An important special case is the wave equation $k = 0$, for which (32) becomes a classical formula covering both even and odd m's. In both cases, for the wave equation it is sufficient to assume that f has not less than $\frac{1}{2}(m + 3)$ derivatives. The assertion made [2, pp. 399*ff.*] that it is sufficient for f to have not less than $\frac{1}{2}(m + 1)$ derivatives and its proof are incorrect.

8. Representation of the solution by the Riemann-Liouville integral. According to a remark of Diaz and Weinberger, by introducing polar coordinates, solution (21), valid for $k > m - 1$, can be written in the form

$$u^{(k)}(x,t) = \frac{2\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k+1-m}{2}\right)\Gamma\left(\frac{m}{2}\right)} \int_0^1 M(x,\rho t;f)(1-\rho^2)^{(k-m-1)/2}\rho^{m-1}\,d\rho, \tag{34}$$

where $M(x,t;f)$ is again a mean value as in (10). Incidentally, the analytic continuation of (34) as a function of k is the basis of the method used by Diaz and Weinberger to obtain a solution for $k < m - 1$ [9].

Let us now introduce the Riemann-Liouville integral

$$I^\beta[F(s)] = \frac{1}{\Gamma(\beta)}\int_0^s (s-\rho)^{\beta-1}F(\rho)\,d\rho, \qquad (\beta > 0). \tag{35}$$

By making the substitution $\sigma = \rho^2 t^2$ in (34) and then putting $t^2 = s$, we easily obtain the formula

$$(36) \qquad u^{(k)}(x,\sqrt{s}) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} s^{(1-k)/2} I^{(k-m+1)/2}[M(x,\sqrt{s};f)s^{(m-2)/2}],$$

which is valid for $k > m - 1$. For $k = m - 1$, (36) reduces to (10).

On the other hand, as u^{k+2n} in (32) is given by a formula of the type (34), the same substitution allows us to write (32) in the form

$$(37) \quad u^k(x,\sqrt{s}) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} s^{(1-k)/2} \frac{d^n}{ds^n} I^{(k+2n-m+1)/2}[M(x,\sqrt{s};f)s^{(m-2)/2}],$$

which is again valid for $k \geq m - 1 - 2n$, $k \neq -1, -3, \cdots$.

9. Connections with the Asgeirsson mean-value theorem. Let us now assume that $k = 0, 1, 2, 3, \cdots$. We shall show first that the EPD equation can be considered as a Coulon, or ultrahyperbolic equation.

Suppose for definiteness's sake that $k < m - 1$. We consider a function $u = u(x_1, \cdots, x_m; y_1, \cdots, y_m)$ which satisfies the Coulon equation

$$(38) \qquad \Delta_x u = \Delta_y u.$$

Then we have the following mean-value theorem [10]:

$$(39) \qquad \frac{1}{\omega_m}\int_{\omega_m} u(x_i + \alpha_i t, 0)\, d\omega_m = \frac{1}{\omega_m}\int_{\omega_m} u(x_i, \alpha_i t)\, d\omega_m.$$

Putting $u(x,0) = f(x)$ and assuming that u depends only on $x_1, \cdots, x_m, y_1, \cdots, y_{k+1}$, we have by the method of descent (Sec. 5)

$$(40) \quad M(x,t;f) = \frac{\omega_{m-k-1}}{\omega_m} \int_{\sum_{i=1}^{m} \alpha_i^2 \leq 1} u(x_i, \alpha_1 t, \cdots, \alpha_{k+1} t) \left(1 - \sum_{i=1}^{k+1} \alpha_i^2\right)^{(m-k-3)/2} d\alpha_1 \cdots d\alpha_{k+1}.$$

If we further assume that u is radially symmetric in the space $y_1, \cdots, y_{k+1}$, and if we put $\rho^2 = \sum_{i=1}^{k+1} \alpha_i^2$, then (40) becomes

$$(41) \qquad M(x_i,t;f) = \frac{\omega_{m-k-1}}{\omega_m}\, \omega_{k+1} \int_0^1 u(x_i, \rho t)(1 - \rho^2)^{(m-k-3)/2} \rho^k\, d\rho.$$

At the same time, equation (38) reduces to the EPD equation,

$$\Delta_x u = u_{tt} + kt^{-1}u_t, \qquad k = 0, \quad 1, \cdots, m-2.$$

Formula (41) has been considered for $k = 0$ by Asgeirsson. As in this special case, it is always an Abel integral equation the solution of which is given by formula (37). The difference between our consideration and Asgeirsson's is that we assume that u in Coulon's equation is radially symmetric in $y_1, \cdots, y_{k+1}$, while Asgeirsson assumes that u depends only on y_1, which he identifies with t. Needless to say, a similar procedure would lead to formula (36) for $k = m-1, m, m+1, \cdots$. However, it should be realized that this procedure does not provide a proof for noninteger values of k and for negative values of k. Even for $k = 0, 1, 2, \cdots$, a verification of the formula would be required to make the application of Asgeirsson's theorem independent of the methods developed in Sec. 7. Moreover, our method does not use Abel's integral equations or analytic continuations.

10. The exceptional cases of the odd negative integers. In the previous sections, the solution has been obtained for $k \neq -1, -3, -5, \cdots$. The only assumption on f was that f should have continuous derivatives of order greater than or equal to $\frac{1}{2}(m+3-k)$. In that case, the solution also has continuous derivatives of the same order with respect to t even for $t = 0$. The situation is quite different in the exceptional cases $k = -1, -3, -5, \cdots$. In these cases, a solution $u^{(k)}$ exists under the same differentiability conditions on f. However, the partial derivative $\partial^{1-k}u^{(k)}/\partial t^{1-k}$ of this solution becomes infinite like $\log t$ at $t = 0$ unless $f(x_1, \cdots, x_m)$ is polyharmonic of order $\frac{1}{2}(1-k)$.

In this generality, a solution was found by Diaz and Weinberger [9] and Blum in an unpublished paper, after the present author pointed out the exceptional role played by polyharmonic initial values. Blum shows that a solution can be obtained directly from (32). B. Friedman was the first to give an example of a solution having nonpolyharmonic initial values, in a communication to the present author.

It is remarkable that for the exceptional values even a small deviation, no matter how regular, from polyharmonic initial values introduces logarithmic infinities in certain derivatives of the solution. In this sense, the solution does not depend continuously on the data.

The exceptional role of polyharmonic initial values can be shown by the following elementary considerations: Let us first take $k = -1$ and assume that $u_{tt}^{(-1)}(x,0)$ exists. By putting $t = 0$ in

$$\Delta u^{(-1)} = u_{tt}^{(-1)} - \frac{u_t^{(-1)}}{t},$$

we find that $\Delta u^{(-1)}(x,0) = 0$, which shows that $f(x)$ must be harmonic.

For $k = -3$, we have

$$u_{tt}^{(-3)}(x,0) = \lim_{t\to 0} \frac{u_t^{(-3)}}{t}.$$

From the EPD equation for $k = -3$, we see that

$$\Delta u^{(-3)}(x,0) = \lim_{t\to 0}\left[-2\frac{u_t^{(-3)}}{t}\right]. \tag{42}$$

From (4) it follows that $t^{-1}u_t^{(-3)} = u^{(-1)}$. If we require $\partial^4 u^{(-3)}/\partial t^4$ to exist for $t = 0$ and that the odd derivatives of $u^{(-3)}$ are zero, then $\partial^2 u^{(-1)}/\partial t^2$ also exists for $t = 0$. Therefore, $\Delta u^{(-1)}(x,0) = 0$ and by (42) we have $\Delta\Delta u^{(-3)}(x,0) = 0$. This remark can be easily generalized to include all the exceptional values. In this case a solution of Cauchy's problem is given by the formula

$$u^{(k)}(x,t) = f(x) + \sum_{h=1}^{-\left(\frac{k+1}{2}\right)} \frac{\Delta^h f}{(k+1)\cdots(k+2h-1)}\frac{t^{2h}}{2\cdot 4\cdots 2h}. \tag{43}$$

For arbitrary $f(x)$, let us give a brief sketch of Blum's method. He observes that in virtue of (3) and (4) it is sufficient to consider the case $k = 1$. To show this, he first puts $s = t^2$, as in Sec. 8, and gets in place of (3) and (4) the relations

$$u_s^{(k)}(x,\sqrt{s}) = u^{(k+2)}(x,\sqrt{s}), \tag{44}$$

$$u^{(k)}(x,\sqrt{s}) = s^{(1-k)/2}u^{2-k}(x,\sqrt{s}). \tag{45}$$

Thus, if $k = -(2r+1)$, repeated application of (44) and (45) yields

$$u^{-(2r+1)} = As^{r+1}u^{2r+3} = As^{r+1}\frac{\partial^{r+1}u^{(1)}}{\partial s^{r+1}}, \tag{46}$$

where A is a constant to be determined later.

It is obvious that if the derivative of $u^{(1)}$ in (46) is finite on $s = 0$, the resulting solution $u^{-(2r+1)}$ will be zero on $s = 0$. In fact, if (46) is to yield a solution $u^{-(2r+1)}$, which assumes nonzero initial values $f(x_1, \cdots, x_m)$, a solution $u^{(1)}$ should be constructed such that $\partial^{r+1}u^{(1)}/\partial s^{r+1}$ has a singularity of the type $s^{-(r+1)}$. This means that $u^{(1)}$ should have a singularity of the type $\log s$. Such a solution is obtained by considering two appropriate solutions $u^{(k)}$ and u^{2-k} for $k = 1 + 2\epsilon$ and then passing to the limit as ϵ approaches zero.

11. Uniqueness. The uniqueness of the solution for $k > 0$ has been established by Asgeirsson by the method of Zaremba. However, for $k < 0$ there can be no uniqueness because any function of the type $t^{1-k}u^{2-k}(x,t)$ which vanishes together with its derivatives with respect to t may be added to a solution $u^{(k)}(x,t)$ of the Cauchy problem.

The situation has been recently clarified by E. K. Blum for the case $m = 1$ in a paper presented to the American Mathematical Society. He showed that, for arbitrary $k < 0$, the difference of any two solutions is of the form $t^{1-k}u^{2-k}(x,t)$, where $u^{2-k}(x,t)$ is a solution such that $\lim_{t\to 0} t^{-k/2}u^{2-k}(x,t) = 0$ and $\lim_{t\to 0} t^{-(k/2)-1}u_t^{2-k}(x,t) = 0$.

12. Conclusion. The key to the solution of all problems considered here was the identities (3) and (4) so long neglected. In conclusion, let us give another application of these identities. We define the Darboux operator, $L_\alpha u = \Delta u - u_{tt} - (\alpha/t)u_t$. Then

$$L_\alpha u^\beta = \frac{\beta - \alpha}{t} u_t^{(\beta)} = u^{\beta+2}$$

for all real α's and β's. Therefore, for any values $\alpha_1, \cdots, \alpha_n$ we have

$$L_{\beta+2n} L_{\alpha_n} \cdots L_{\alpha_2} L_{\alpha_1} u^{(\beta)} = 0.$$

Taking $\beta = 0$ and $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 2n$, we obtain

$$L_{2n}^{n+1} u^0(x_1, \cdots, x_m, t) = 0.$$

Putting $m = 1$, we obtain immediately a theorem of Friedrichs [2, pp. 416*ff*.].

Bibliography

1. G. Darboux, *Leçons sur la théorie générale des surfaces*, Vol. II, Paris 2d ed., 1915 Chaps. III, IV.
2. R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, Vol. II, Berlin, 1937.
3. M. H. Martin, *Riemann's method and the problem of Cauchy*, Bull. Amer. Math. Soc. vol. 57 (1951) pp. 238–249.
4. M. H. Martin and J. B. Diaz, *Riemann's method and the problem of Cauchy*, II, *The wave equation in n dimensions*, Proc. Amer. Math. Soc. vol. 3 (1952) pp. 476–484.
5. M. B. Kapilevic, *On an equation of mixed elliptic-hyperbolic type*, Rec. Math. (Mat. Sbornik) N.S. vol. 30 no. 72 (1952) pp. 11–38.
6. A. Weinstein, *Sur le problème de Cauchy pour l'équation de Poisson et l'équation des ondes*, C.R. Acad. Sci., Paris vol. 234 (1952) pp. 2584–2585.
7. J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*, Yale University Press, New Haven, 1923.
8. B. B. Baker and E. T. Copson, *The mathematical theory of Huyghens' principle*, Oxford, 2d ed., 1950.
9. J. B. Diaz and H. F. Weinberger, *A solution of the singular initial value problem for the Euler-Poisson-Darboux equation*, Proc. Amer. Math. Soc. vol. 4 (1953) pp. 703–715.
10. L. Asgeirsson, *Uber eine Mittelwerteigenschaft von Losungen homogener linearer partiellen Differentialgleichungen* 2, *Ordnung mit konstanten Koeffizienten*, Math. Ann. vol. 113 (1937) pp. 321–346.

Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Md.

ON LIÉNARD'S DIFFERENTIAL EQUATION

BY

S. LEFSCHETZ

The following differential equation

$$\ddot{x} + \frac{dF(x)}{dx}\dot{x} + g(x) = 0 \tag{1}$$

has been investigated by Liénard. He proved that under certain rather general conditions it possesses a unique periodic solution. A special case of (1) is the famous equation of van der Pol, which corresponds to $F = \mu(x^3/3 - x)$, $g(x) = x$. We propose to discuss an intermediary case, *viz.*, the equation

$$\ddot{x} + \frac{dF(x)}{dx}\dot{x} + x = 0. \tag{2}$$

We shall assume that: (*a*) $F(x)$ is an odd analytic function of x for all real x's and has only a finite number of zeros; (*b*) $y = F(x)$ behaves like a "vertical" parabola $y = x^n$, $n \geqq 3$; (*c*) from a certain finite positive value of x on, $F(x)$ is monotone increasing with the obvious performance caused by (*a*) for x large negative.

As usual, we introduce the associated pair of equations

$$\dot{x} = y - F(x), \qquad \dot{y} = -x, \tag{3}$$

and conditions (*a*), (*b*), (*c*) have for consequence (from Liénard's treatment) that (3) has a finite number of limit cycles surrounding the origin. The outer limit cycle is stable outside. We have recently discussed the pattern of the solutions of van der Pol's system outside the limit cycle (see a forthcoming paper in a monograph of the *Annals of Mathematics Studies*). We propose to extend the same results to the system (3). That is to say, we propose to show that all the paths outside the outer limit cycle spiral clockwise around that limit cycle.

Basically our study rested upon the behavior of the paths "at infinity." Now the appropriate mechanism, which goes back in substance to Poincaré, did consist in extending (3), or rather the equivalent equation

$$x\,dx + (y - F(x))\,dy = 0, \tag{4}$$

to the whole projective plane and examining the new critical points "at infinity." That is to say, one replaces everywhere x by x/z, y by y/z, thus homogenizing the system. A point of the extended plane is now a triple (x,y,z) of numbers not all zero. Moreover (x,y,z) and (kx,ky,kz) represent the same point. The old points correspond to $z \neq 0$, and one may assume for them that $z = 1$; hence $dz = 0$. New points, however, are introduced, *viz.*, on the line $z = 0$, and they correspond to the directions on the xy-plane.

Now for a system

$$\dot{x} = P(x,y), \qquad \dot{y} = Q(x,y),$$

where P and Q are polynomials, I have shown that the extension is given by

$$\begin{vmatrix} dx & dy & dz \\ x & y & z \\ z^n P\left(\frac{x}{z}, \frac{y}{z}\right) & z^n Q\left(\frac{x}{z}, \frac{y}{z}\right) & 0 \end{vmatrix} = 0,$$

where n is the largest of the degrees of P and Q. The study of this relation along the line $z = 0$ proceeds practically as for the initial system. I have found that with a few minor modifications the same thing holds for the system (3). One verifies readily that the equation

$$\text{(5)} \qquad \begin{vmatrix} dx & dy & dz \\ x & y & z \\ y - zF\left(\frac{x}{z}\right) & -x & 0 \end{vmatrix} = 0$$

reduces to (4) upon making $z = 1$, $dz = 0$, and so it represents the appropriate extension of (4). The behavior of the paths of (3) at infinity is pretty largely governed by the behavior of the solutions of (5) about the critical points on $z = 0$.

Now it is found that, as in the case of van der Pol's equation, the only critical points of (5) on $z = 0$ are point A at infinity on the x-axis, or point $(1,0,0)$, and point B at infinity on the y-axis, or point $(0,1,0)$.

Behavior at point A. We may take $x = 1$, $dx = 0$. The expanded equation becomes

$$z\left[y - zF\left(\frac{1}{z}\right)\right] dy - \left[1 + y\left(y - zF\left(\frac{1}{z}\right)\right)\right] dz = 0.$$

We are interested only in values of y,z near zero. This makes

$$zF\left(\frac{1}{z}\right) = \frac{F(1/z)}{1/z}$$

very large. Upon dividing by this expression, the terms which will have it in the denominator may be neglected, and what is left is, dropping the factor z, $z\,dy = y\,dz$. This integrates to $y = Cz$ and shows that, as in the van der Pol case, A is the simplest type of node: the paths stream away from it in all directions.

Behavior at point B. The full analysis of this point merely requires to prove the following property: there is just one path other than $z = 0$ tending to that point in the first quadrant. Now near B we may make $y = 1$, $dy = 0$. The equation thus becomes

$$zx\,dx = \left[x^2 - zF\left(\frac{x}{z}\right) + 1\right] dz,$$

which may be written

$$\frac{dx}{dz} = \frac{x}{z} - \frac{1}{x}F\left(\frac{x}{z}\right) + \frac{1}{xz} \tag{6}$$

or again

$$\frac{dx}{d \log z} = x - \frac{F(x/z)}{x/z} + \frac{1}{x}.$$

An examination of the solutions shows that they are parabolic, and in the xy-plane above the curve $y = F(x)$. Let us suppose that we have two paths γ_1, γ_2 tending to B and that γ_1 is above γ_2. Then $x/z \to +\infty$ along these paths, and $F(x/z)/(x/z)$ is at least as large as $(x/z)^2$. Let $x_1(z)$, $x_2(z)$ be the two solutions corresponding to γ_1 and γ_2, and set $x_1 - x_2 = u > 0$. Then from (6)

$$\frac{du}{d \log z} = u - \frac{u}{x_1 x_2} - [\varphi(m_1) - \varphi(m_2)], \qquad \varphi(m) = \frac{F(m)}{m}, \qquad m = \frac{x}{z}. \tag{7}$$

Now along γ_1, γ_2 almost

$$\varphi(m_1) = C_1 x^{p_1}, \qquad \varphi(m_2) = C_2 x^{p_2},$$

where either $p_1 > p_2$ or else $p_1 = p_2$ and $C_1 > C_2$. Hence in (7) at the right the first term is small for z small, the second and third are negative, and the third is large. Hence for z small $du/d \log z < 0$. Thus u increases as $z \to 0$ contrary to the assumption that γ_1, γ_2 both $\to B$. Hence only one such path can exist.

We have therefore proved our assertion that all paths outside the last limit cycle tend spirally to it. The signs of $\dot{x}, \dot{y}$ show then that the spiraling is clockwise.

We shall now examine another type of question closely related to the result just obtained. Consider the motion under a forced oscillation represented by

$$\ddot{x} + \frac{dF(x)}{dx}\dot{x} + x = e(t), \tag{8}$$

where e is continuous and periodic with period τ (hence bounded) and $F(x)$ is as before. The question is to prove that there is a solution of (8) which is likewise periodic and of period τ. The problem is evidently equivalent to finding a closed trajectory of

$$\dot{x} = y - F(x), \qquad \dot{y} = -x + e(t) \tag{9}$$

described in the time τ. By a well-known procedure based on Brouwer's fixed-point theorem, the problem reduces to finding a simple closed curve Γ at all points of which the vector W, whose components are the right-hand sides of (9), points inward. Now it so happens that, in the literature, the generality of the system analogous to (9) causes Γ to be quite complicated. We propose to show that in the present instance Γ may be chosen exceptionally simple.

Let generally $|V|, \cdots$, denote the magnitude of the vector $V, \cdots$. Let more particularly V be the vector whose components are the right-hand

sides of (3). At points $M(x,y)$ arbitrarily far from the origin, $|V|$ is arbitrarily large. For

$$|V|^2 = (y - F(x))^2 + x^2,$$

and so either x is large, and $|V|$ likewise, or else x is bounded and y is large, hence $|V|$ is again large. Since $|V - W| = |e(t)|$ is bounded, to show that W points inward along Γ, it is sufficient to show that Γ is sufficiently far out and that V points inward (never tangentially) along Γ.

To construct an appropriate Γ, we exploit an idea which the present author developed in his Princeton seminar of three years ago. It rests upon the consideration of the system

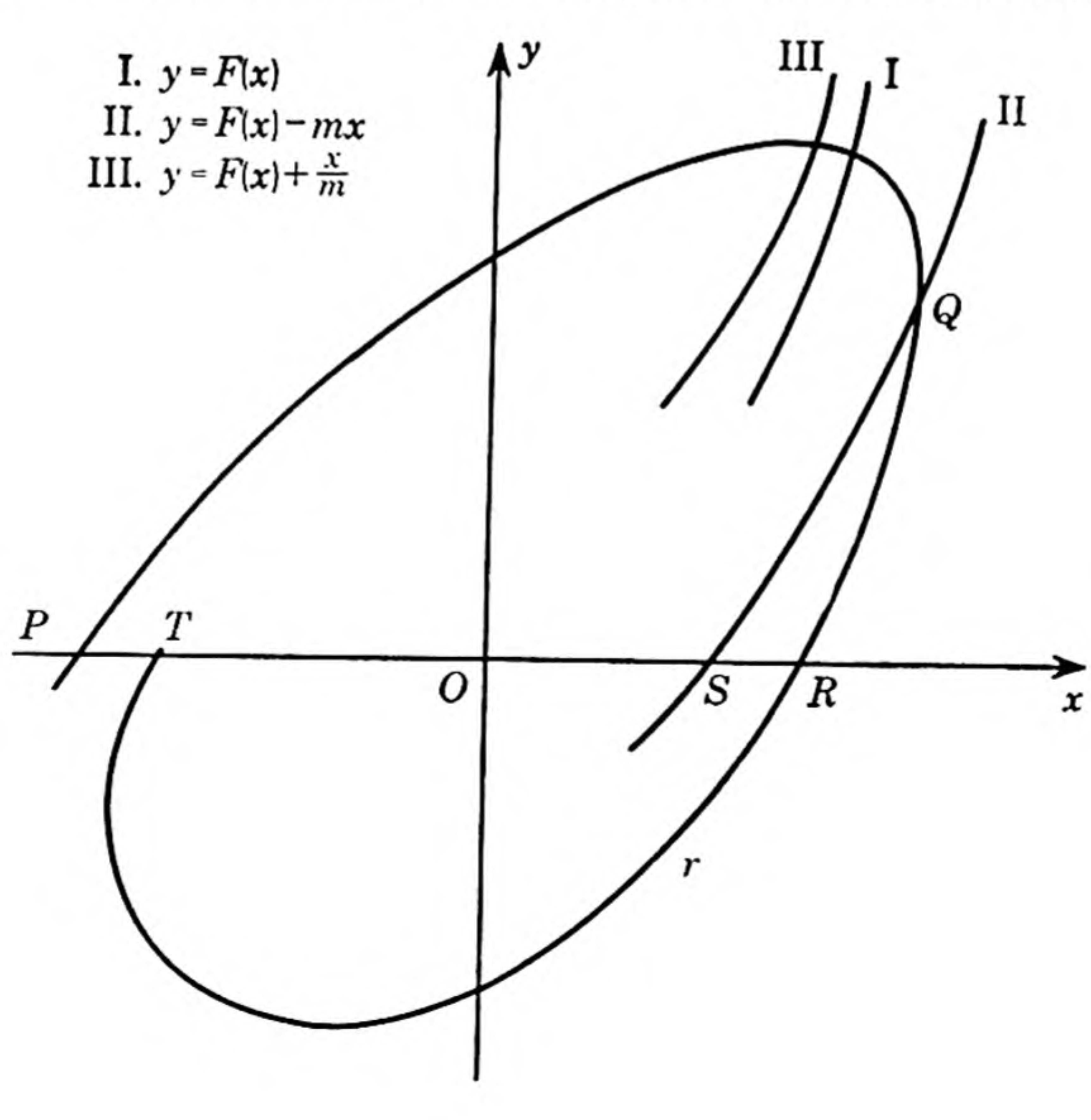
$$(10) \qquad \begin{cases} \dot{x} = (y - F(x)) \cos \alpha + x \sin \alpha, \\ \dot{y} = (y - F(x)) \sin \alpha - x \cos \alpha, \\ 0 \leqq \alpha < \frac{\pi}{2}, \qquad \tan \alpha = m, \end{cases}$$

whose vector V_α is merely V rotated counterclockwise by the fixed angle α. The same method as before shows that (10) has perhaps a singular point at infinity at B (on the y-axis) and certainly at the point A_α at infinity on the line $y = mx$. Since there are no asymptotes in directions in the second and fourth quadrants, all the solutions of (10) far out cross the x-axis. The curves

$$\text{(II)} \qquad y = F(x) - mx,$$

$$\text{(III)} \qquad y = F(x) + \frac{x}{m}$$

delimitate the regions $\dot{x} \geqq 0$, $\dot{y} \geqq 0$. Since far out they are in the first quadrant, a solution of (10) starting at a point P on the x-axis far to the left behaves as in the figure and is arbitrarily far out, if P is remote enough. That is to



say, the arc PQ is arbitrarily far out if P is sufficiently remote. Now R is to the right of S where OS is the largest root of $F(x) = mx$, or the largest x of an intersection of $y = F(x)$ with the line $y = mx$. Thus for m sufficiently large, OS, hence OR, may be made arbitrarily large. Moreover increasing m puts R even farther to the right. As a consequence, for m and OP large enough the point R, and hence by symmetry the arc RT, will be arbitrarily far out. Hence the full solution of (10) from P to T will be arbitrarily far out. Furthermore T is to the right of P, since otherwise the path $PQRT$ extended below P would cross itself, which is ruled out. The arc $PQRT$ of the solution of (10) plus the segment PT constitutes Γ. By construction along $PQRT$ the vector V makes with the tangent to the arc directed forward a clockwise angle α and hence points inward throughout. Along the open segment PT the vector V points above the segment and hence again inward relative to Γ. Since Γ is arbitrarily far out, it satisfies all the required conditions, to establish that (8) does have a periodic solution of period τ.

It may be observed that all the considerations of the present paper apply with only insignificant modifications if (1) is replaced, say by the more general Liénard equation,

$$\ddot{x} + \frac{dF(x)}{dx}\dot{x} + g(x) = 0,$$

where $g(x)$ behaves like $F(x)$ and for x very large the order of $g(x)$ is less than that of $F(x)$. One will also replace (8) by

$$\ddot{x} + \frac{dF(x)}{dx}\dot{x} + g(x) = e(t),$$

where $e(t)$ is as before.

PRINCETON UNIVERSITY,
PRINCETON, N.J.

THE EFFECT OF SMALL CONSTRAINTS ON NATURAL VIBRATIONS[1]

BY

R. J. DUFFIN AND A. SCHILD

1. Introduction. The addition of a constraint to a vibrating linear mechanical system tends to increase the frequency of each normal mode. It is the aim of this paper to show how this qualitative statement can be expressed in a quantitative form. Consider constraints which can be varied continuously. We wish to find the small increase in frequency resulting from a small increase of constraint.

The original system will be called the *unconstrained* system S; the system with the small additional constraints, the *constrained* system S'. Our problem is resolved by the following simple formula:

$$(A) \qquad \delta\lambda = \delta V.$$

Here λ is the square of the angular frequency of a given normal mode of vibration of S, and $\delta\lambda = \lambda' - \lambda$ is the increase in λ resulting from the increase in constraints. To maintain this increase in constraint, external forces of amplitude f must be applied to S; the resulting vibration of S' is assumed to be suitably normalized. Now let the unconstrained system S be at rest. Let the forces f act to produce a static deflection. Then δV is the potential energy of S in this state of static deflection.

In many cases the normalized mode of vibration is known for the constrained system S'. The forces f can then be deduced. Hence $\delta\lambda$ can be obtained without knowing any of the other constrained or unconstrained modes.

The proof of formula (A) can be carried out abstractly using the methods of Hilbert space. The essential hypothesis needed about S is that the kinetic-energy quadratic form T be completely continuous relative to the potential-energy quadratic form V. The complete proof will be given in another paper. Here applications to plates, membranes, and other systems are sketched.

Formula (A) reduces the dynamic problem to a problem in statics. To facilitate the solution of this static problem, the following formula is employed:

$$(B) \qquad \delta V = \delta V^*.$$

Let S^* be a system which coincides with S in the neighborhood of the applied forces f of additional constraint. Outside of this neighborhood, S^* is essentially arbitrary. Formula (B) asserts that, under the same forces f, the potential energies acquired by S and S^* are equal. An equivalent way of stating

[1] This paper is based on research conducted in part under a contract between the Flight Research Laboratory of the U.S. Air Force and Carnegie Institute of Technology, Contract No. AF-33(616)-294.

(B) is as follows: Close to a rigid constraint, the stiffness of an elastic body S is essentially independent of other constraints which are far away. Formula (B) is a general statement similar in nature to St. Venant's principle. In the applications S^* is taken to be an infinite system because this simplifies the analysis.

There are many applications of the basic problem considered. Various special cases have been treated in the literature by methods somewhat different from those proposed here. It appears that these treatments lack unity and rigor.

2. The clamped plate. The motivation for the present investigation was a nodal-line problem in the vibration of clamped plates. Consider a thin plate which is clamped on a closed curve B bounding a plane region R. If q is a deflection of the plate perpendicular to its surface, then in R the relation satisfied by a normal mode of vibration is

$$\Delta\Delta q = \lambda q. \tag{1}$$

Here Δ is the Laplacian operator, and the eigenvalue λ is the square of the angular frequency of vibration. On the boundary B the clamping conditions are

$$q = \frac{\partial q}{\partial n} = 0. \tag{2}$$

An important result in the theory of vibrations states that nodal lines cannot occur in the vibration of lowest frequency of a membrane. This was shown by Rayleigh. The question has been raised by Weinstein whether or not the same statement is true for the clamped plate. In other words, for the smallest value of λ, is it possible that q will change sign in R? Weinstein's question is related to several other problems in vibrations. The one-dimensional analogue of the plate is the beam: it was shown by Rayleigh that the vibration of lowest frequency of a beam clamped at both ends cannot have a node. In a recent paper, Szegö [3] has shown that, for clamped plates of a given area, the circular plate has the lowest frequency of vibration; he derives this result under the explicit assumption that nodal lines are not present.

Weinstein's question has been investigated from three different lines of approach. In the first a semi-infinite strip is considered; its behavior suggests that there will be nodal lines in the lowest mode of vibration of a clamped plate in the shape of a long rectangle overlapping a circle of suitable radius [1]. In the second approach the doubly connected region between two concentric circles is considered; this eigenvalue problem can be solved explicitly and, for a sufficiently small inner circle, the lowest mode of vibration has a node along a diameter [2]. The third approach constitutes the main subject of this paper. This approach is more general and has applications to many problems of vibrations besides that of the plate.

Let R be a region, and let $\bar{R}$ be its mirror image in a tangent line l, which we choose as our x-axis (Fig. 1). Imagine a membrane or plate constrained on the boundary of $R + \bar{R}$. Although there are actually two distinct regions, we can consider the eigenvalue problem for the whole region $R + \bar{R}$ simultaneously. Corresponding to each eigenvalue λ' there will be two eigenfunctions. One eigenfunction (q'^{+}) may be taken to be even with respect to the line l, the other (q'^{-}) odd. Then

$$
\begin{aligned}
q'^{+}(x,y) &= q'^{+}(x,-y), \\
q'^{-}(x,y) &= -q'^{-}(x,-y).
\end{aligned}
\tag{3}
$$

Let us now open a small gap of length 2ϵ between the two regions (Fig. 2). We shall call the eigenvalue problem with the gap open the unconstrained

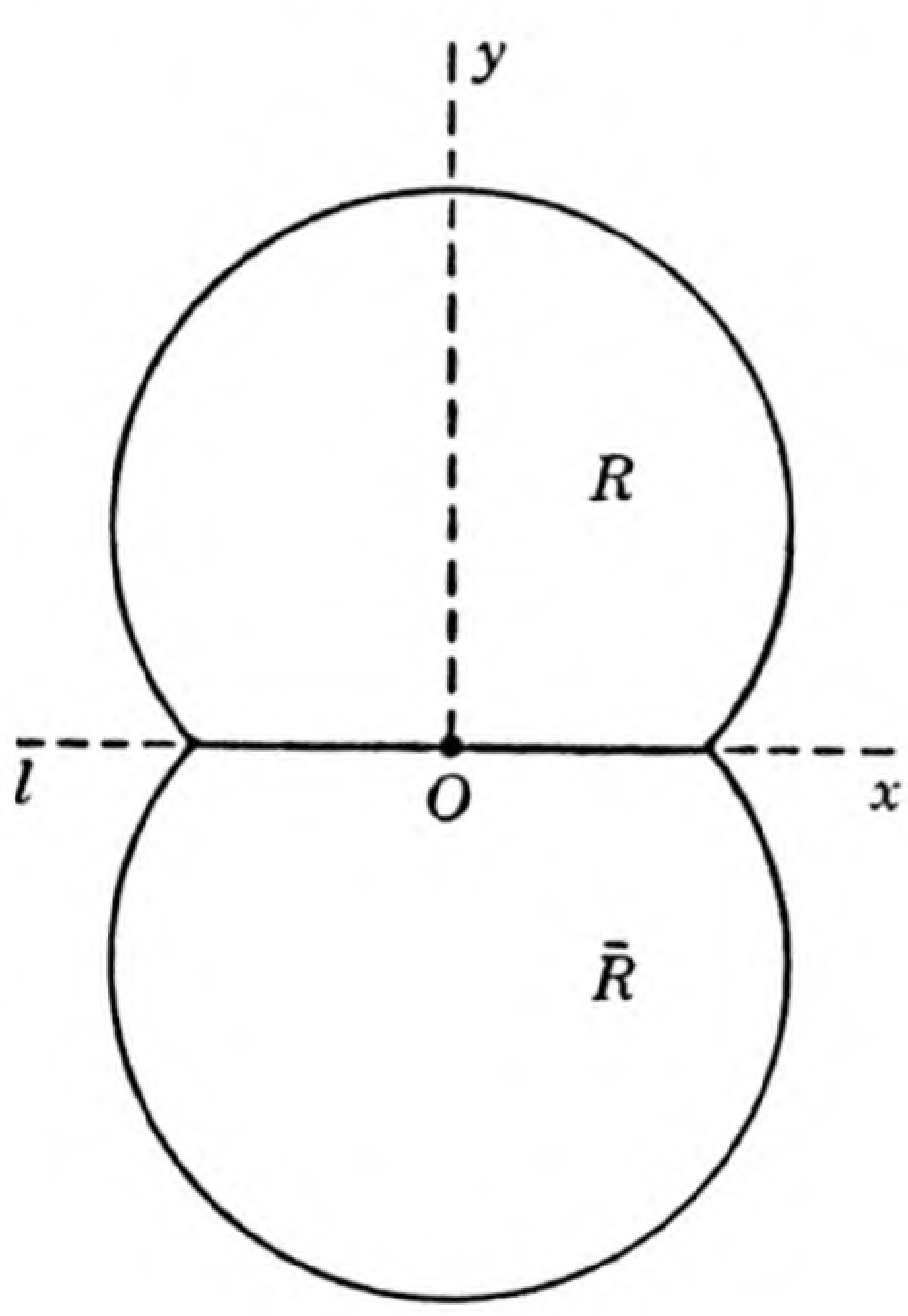


FIG. 1

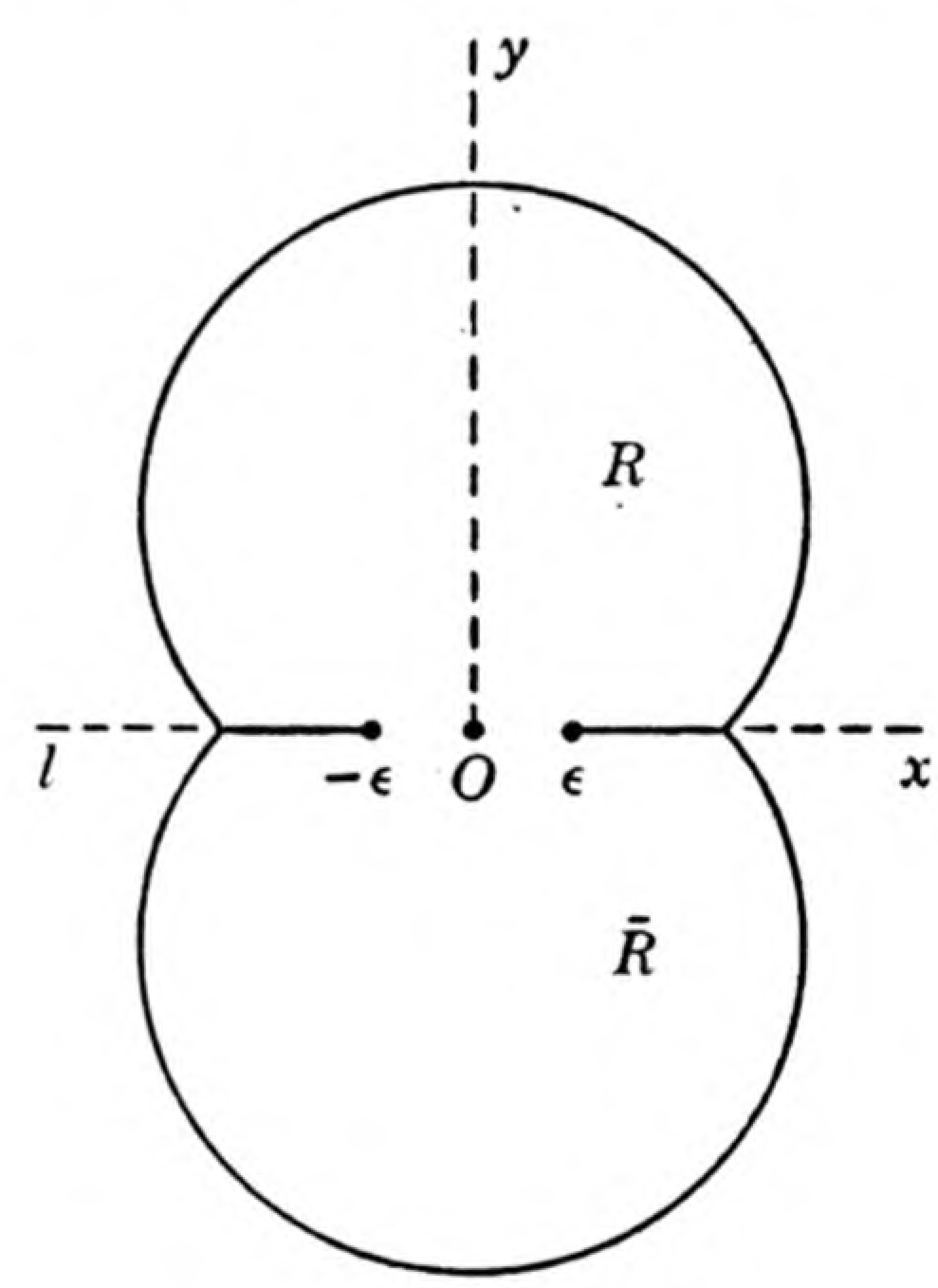


FIG. 2

problem. Again the eigenfunctions will be even (q^{+}) or odd (q^{-}). But, in general, these will now have different eigenvalues. The constrained problem of Fig. 1 is degenerate, each eigenvalue λ' corresponding to at least two eigenfunctions. It is to be expected that opening the gap removes this degeneracy, so that each eigenvalue λ' splits into two, λ^{+} corresponding to even states, λ^{-} corresponding to odd states. This splitting of degenerate eigenvalues is, of course, a very common phenomenon. The Zeeman and Stark effects in quantum mechanics are well-known examples.

The general formula (A) gives a first-order perturbation expression for the spectral fine structure. In the case of the clamped plate we shall see in Sec. 5 that λ^{-} is less than λ^{+} for small ϵ. It follows that the mode of lowest frequency must be odd. The odd function has a nodal line at the gap. This is another counter-example which answers Weinstein's question.

3. General systems. We shall now give a short sketch of the method used to derive formula (A). The details of the proof will be published later.

Consider the small vibrations of a general mechanical system S whose configuration is given by a vector q in a Hilbert space $\mathfrak{H}$. We can analyze such

a system in terms of a potential-energy quadratic form $V(q,q)$ and a kinetic-energy quadratic form $T(q,q)$. The amplitude q of a normal mode satisfies

$$V(q,x) - \lambda T(q,x) = 0 \tag{4}$$

for all vectors x in $\mathfrak{H}$. It is customary to normalize the eigenfunctions so that

$$T(q,q) = 1. \tag{5}$$

Then it follows that

$$\lambda = V(q,q). \tag{6}$$

In fact, λ may be defined as a stationary value of $V(q,q)$ subject to (5). The stationary vector q satisfies the Eulerian equation (4).

We now imagine that the system S is prevented from moving in a certain small region in the neighborhood of already existing constraints. This problem with the small additional constraints, the constrained problem S', is characterized by equations similar to (4) and (5):

$$V(q',x') - \lambda' T(q',x') = 0, \tag{7}$$
$$T(q',q') = 1. \tag{8}$$

However, the vectors q', x' are restricted by the condition that they vanish in a subspace $\mathfrak{F}$ of $\mathfrak{H}$; using physical language, $\mathfrak{F}$ will be called the region of constraint. More precisely, if $\mathfrak{H}'$ is the orthogonal complement of $\mathfrak{F}$ in $\mathfrak{H}$ with the potential-energy form as metric,

$$V(\mathfrak{F},\mathfrak{H}') = 0, \tag{9}$$

then we demand that q', x' be vectors in $\mathfrak{H}'$. In a certain sense, which we shall not specify here, the subspace $\mathfrak{F}$ must be small, corresponding to the physical fact that the additional constraints are small.

For x in $\mathfrak{H}$, the expression $V(q',x) - \lambda' T(q',x)$ is a linear functional of x which vanishes in $\mathfrak{H}'$. Therefore, there must exist a vector ϕ such that

$$V(q',x) - \lambda' T(q',x) = V(\phi,x), \tag{10}$$

and ϕ must be in $\mathfrak{F}$, that is,

$$V(\phi,x') = 0. \tag{11}$$

The configuration of the system described by ϕ is precisely that of the static deflection discussed in the Introduction. It is the static deflection of the unconstrained system S under forces f which are equal to the amplitudes of the additional forces of constraint acting in the vibration of S' when this vibration is normalized by (8).

Subtracting (10) with $x = q$ from (4) with $x = q'$, we obtain

$$\begin{aligned}(\lambda' - \lambda)T(q,q') &= -V(\phi,q) \\ &= V(\phi,\phi) - V(\phi,\phi + q).\end{aligned} \tag{12}$$

For small constraints q and q' are close so that $T(q,q')$ is essentially unity. Also, it can be shown that the term $V(\phi,\phi + q)$ is of smaller order of magnitude than $V(\phi,\phi)$. Roughly speaking, this is due to the fact that in the region of additional constraints the kinetic energy is less important than the potential energy so that in $\mathfrak{F}$ the dynamic deflection q and the static deflection ϕ are essentially equal in magnitude. Our sign conventions are such that $\phi + q$ is then essentially zero in $\mathfrak{F}$ and hence $V(\phi,\phi + q)$ is small compared with $V(\phi,\phi)$. Ignoring the small terms, (12) becomes

$$\lambda' - \lambda = V(\phi,\phi), \tag{13}$$

and this is, of course, our formula

$$\delta\lambda = \delta V. \tag{A}$$

Formula (A) holds in the simple form given here when the eigenvalue λ' is nondegenerate. In the case of degenerate states the frequency changes are obtained by solving a secular equation.

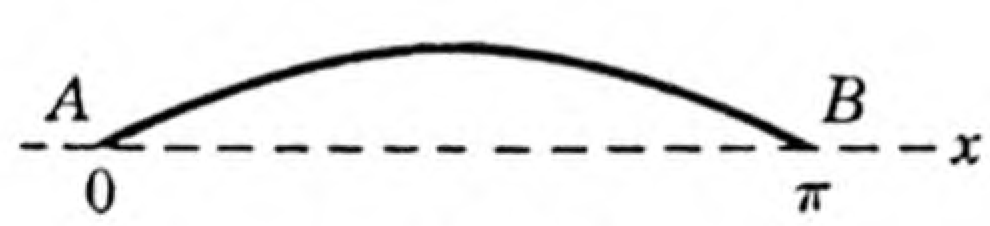


FIG. 3

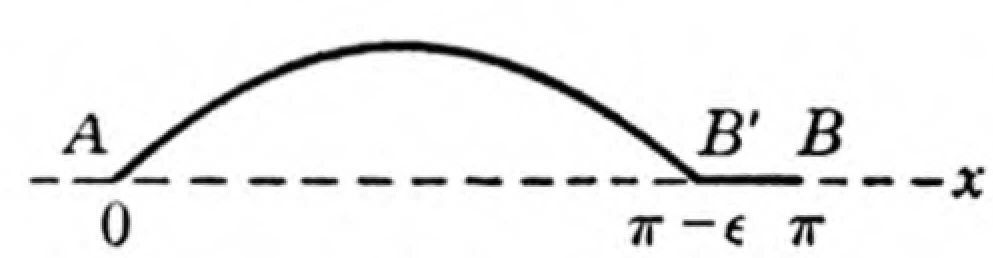


FIG. 4

Formula (A) is also valid with a dual interpretation of the potential energy δV in terms of the displacements of the original system in the region of the additional constraints. In this formulation, $\delta\lambda$ can be obtained from the knowledge of a mode of vibration of the unconstrained system.

4. The string. To illustrate our formula, we shall apply it to a simple problem where the solutions are known and where the approximate change in frequency given by (A) can be compared with the rigorous value. Consider a violin string stretched between points A at $x = 0$ and B at $x = \pi$ (Fig. 3). The violinist presses the string with his finger at a distance ϵ from one end (Fig. 4). This raises the tone. For a suitable choice of the physical units, the normalized nth mode of the constrained system AB' is given by

$$q'_n = 2(\pi - \epsilon)^{-\frac{1}{2}} \sin \frac{n\pi}{\pi - \epsilon} x, \qquad \frac{1}{2}\int_0^{\pi-\epsilon} q'^2_n \, dx = 1. \tag{14}$$

The force of constraint at B' is

$$f_n = \left| \left[\frac{dq'_n}{dx} \right]_{x=\pi-\epsilon} \right| = 2n\pi(\pi - \epsilon)^{-\frac{3}{2}}. \tag{15}$$

Under a force f_n at B' the static deflection ϕ_n of the unconstrained string AB is as shown in Fig. 5. The deflection at B' is

$$\phi_n(B') = \epsilon\pi^{-1}(\pi - \epsilon)f_n. \tag{16}$$

The potential energy is therefore

$$\text{(17)} \qquad \begin{aligned} \delta V_n &= \tfrac{1}{2}\phi_n(B')f_n \\ &= 2\epsilon\pi^{-1}n^2 + O(\epsilon^2). \end{aligned}$$

Formula (A) states that, correct to the lowest order in ϵ, the change in the square of the angular frequency is

$$\text{(18)} \qquad \delta\lambda_n = 2\epsilon\pi^{-1}n^2.$$

This result agrees with the known rigorous expression

$$\text{(19)} \qquad \delta\lambda_n = n^2\left(1 - \frac{\epsilon}{\pi}\right)^{-2} - n^2 = 2\epsilon\pi^{-1}n^2 + O(\epsilon^2).$$

To show how this calculation may be simplified by application of formula (B), consider the system S^* in which the point A has been allowed to recede

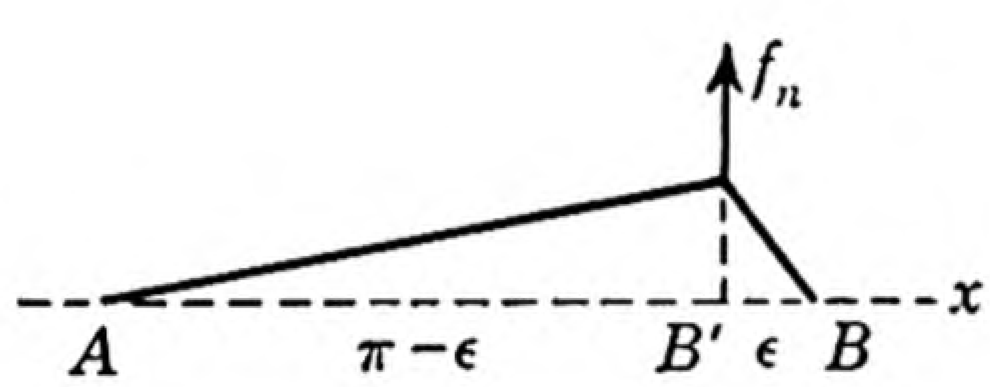


Fig. 5

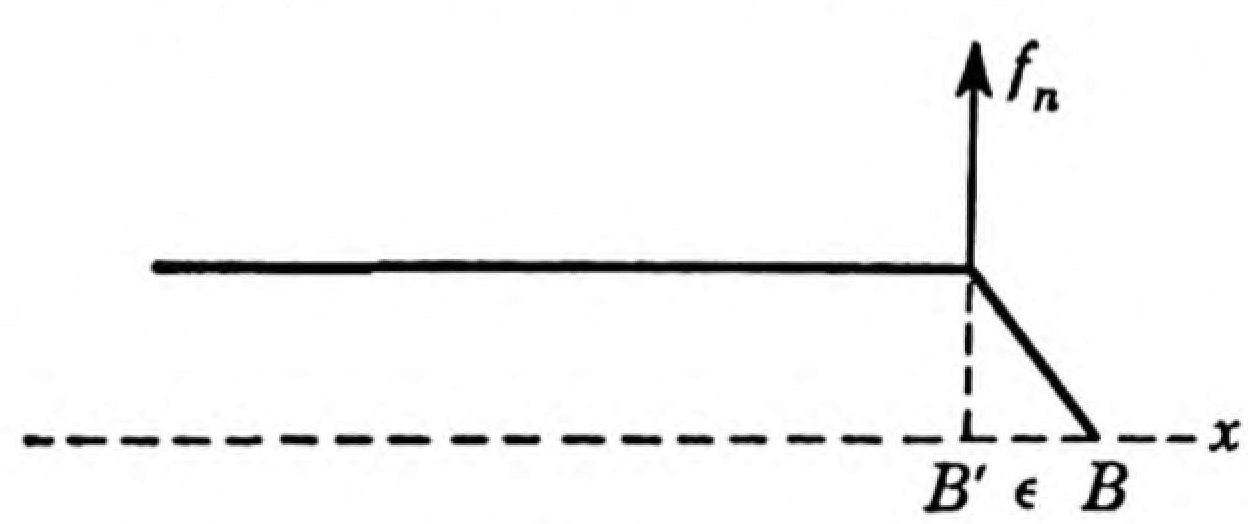


Fig. 6

to $x = -\infty$. Under the force f_n at B' the static deflection ϕ_n^* is as shown in Fig. 6, and

$$\text{(20)} \qquad \phi_n^*(B') = \epsilon f_n = \phi_n(B') + O(\epsilon^2).$$

Therefore

$$\text{(21)} \qquad \delta V^* = \delta V + O(\epsilon^2).$$

The problem of the vibrating string is less trivial if the string has a variable distribution of mass along its length. The present methods still apply.

5. The clamped plate (Continued). In the clamped-plate problem of Fig. 1 we shall now make the simplifying assumption that the individual region R is sufficiently asymmetric so that its normal vibrations are nondegenerate. The problem of the combined region $R + \bar{R}$ is then twofold degenerate, but this degeneracy is due only to the symmetry of reflection in the line l, and this symmetry is preserved when the gap 2ϵ is opened. The even functions q^+ of the unconstrained problem tend to the even functions q'^+ of the constrained problem as $\epsilon \to 0$, and similarly for the odd functions. Therefore, the even and odd vibrations can be treated independently, and formula (A) can be applied to each; this avoids a discussion of the secular equation mentioned near the end of Sec. 3.

In the nth even vibration of the constrained plate the constraining force along the gap consists of a linear force density $|\partial^3 q_n'/\partial y^3|$. In the nth odd vibration the constraining force along the gap consists of a linear density

$|\partial^2 q_n'/\partial y^2|$ of force moments. The functions q_n' describe the deflection of the single region R when the gap is closed. Therefore, q_n' and its derivatives are continuous along the gap and, since ϵ is small, we can replace the force or moment densities of the constraining forces by their constant values at the center O of the gap:

$$
(22)\qquad
\begin{aligned}
f_n &= \left|\frac{\partial^3 q_n'}{\partial y^3}\right|_0, \\
m_n &= \left|\frac{\partial^2 q_n'}{\partial y^2}\right|_0.
\end{aligned}
$$

Let ϕ_n^+ and ϕ_n^- be the static deflections of the unconstrained plate under these respective force distributions along the gap; ϕ_n^+ will be an even deflection, ϕ_n^- odd. Then the potential energy δV_n is given by

$$
(23)\qquad
\begin{aligned}
\delta V_n^+ &= \frac{1}{2} f_n \int_{-\epsilon}^{\epsilon} \phi_n^+(x,0)\, dx, \\
\delta V_n^- &= \frac{1}{2} m_n \int_{-\epsilon}^{\epsilon} \left[\frac{\partial \phi_n^-}{\partial y}\right]_{y=0} dx.
\end{aligned}
$$

Thus δV is determined by the behavior of the function ϕ in the immediate neighborhood of the gap. Formula (B) states that, to the lowest order in ϵ, this potential energy will depend only on the part of the boundary at distances of order ϵ from the gap and not on the shape of the boundary at finite distances from the gap.

To make the argument more convincing, let us take the following example: Consider a large circular plate clamped around the rim. Here the Green's function is known, and it can be verified that near the center of the plate the stiffness is low. This means that a small force acting there produces a large displacement; a small couple distributed along a unit segment near the center produces a large angular tilt. Let us now impose additional clamping along a diameter of the circle except for a finite gap at the center (Fig. 7). Now the plate is stiff in the neighborhood of the center: a small force acting there produces a small displacement, and a small couple distributed along a unit segment produces a small tilt. This shows that the stiffness near the gap is essentially determined by the clamping along the diameter and not by the boundary conditions on the distant rim.

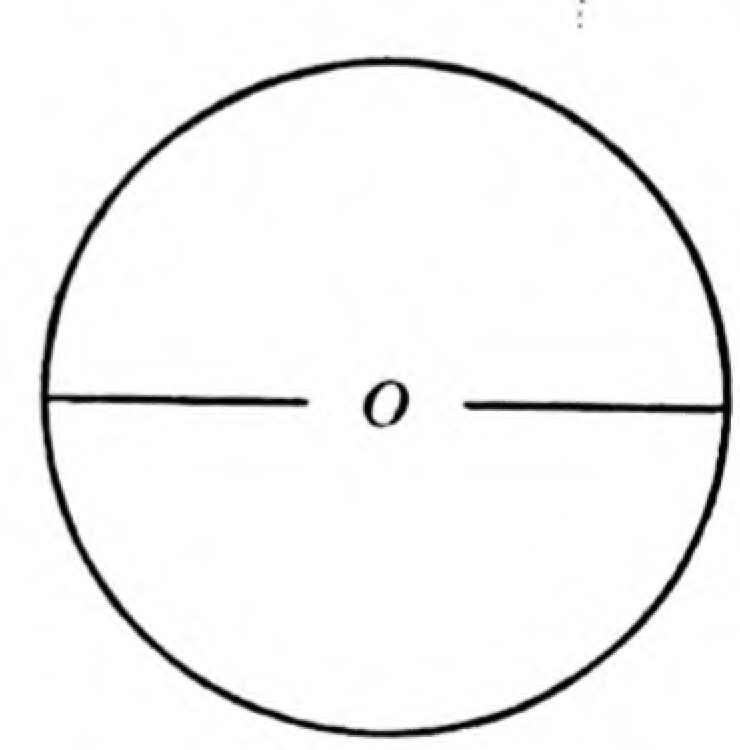


FIG. 7

This physical discussion suggests that the principal part δV^* of the potential energy δV can be determined from the static deflection of an infinite plate clamped along two half lines which are separated by a gap 2ϵ as shown in Fig. 8. With this simplifying approximation the frequency increments can now be obtained explicitly, except for two constants which are independent of the geometry of our problem. Consider the standardized problem of an infinite plate clamped along two half lines which are separated by a gap of

length 2 as shown in Fig. 9. Let v^+ be the potential energy under static deflection produced by a unit force density acting across the gap, and let v^-

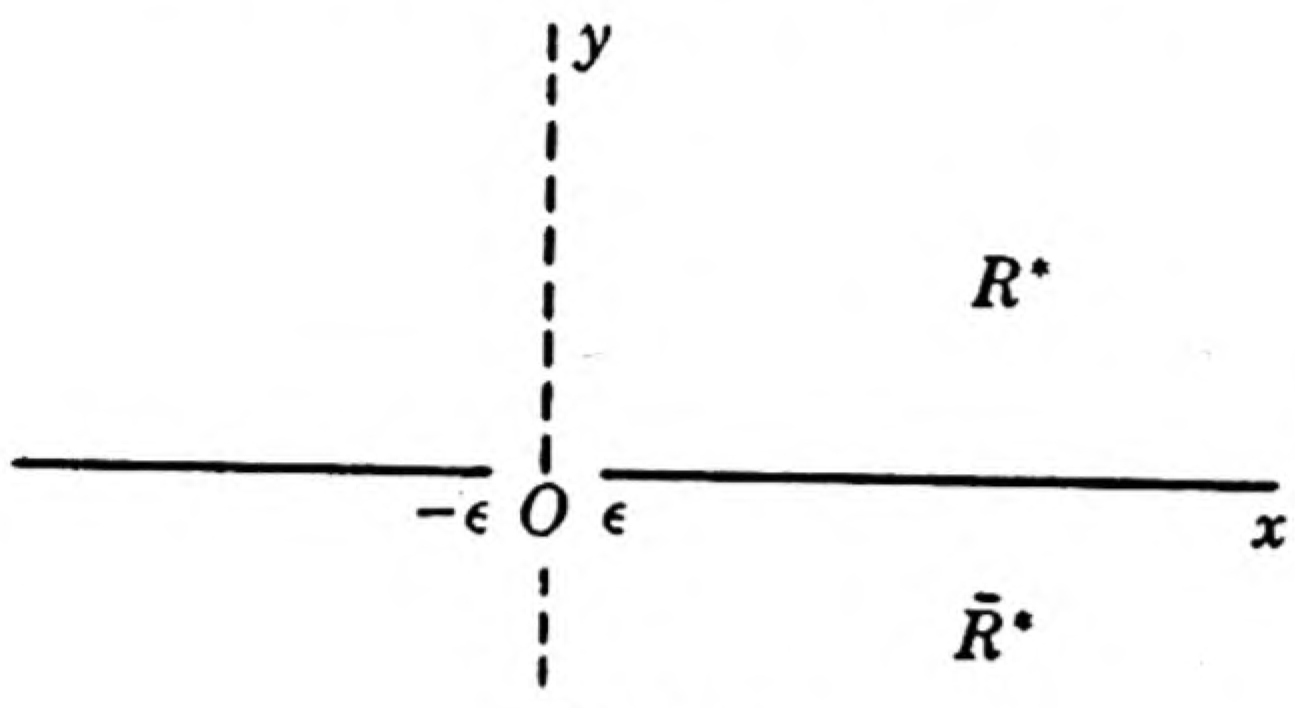


FIG. 8

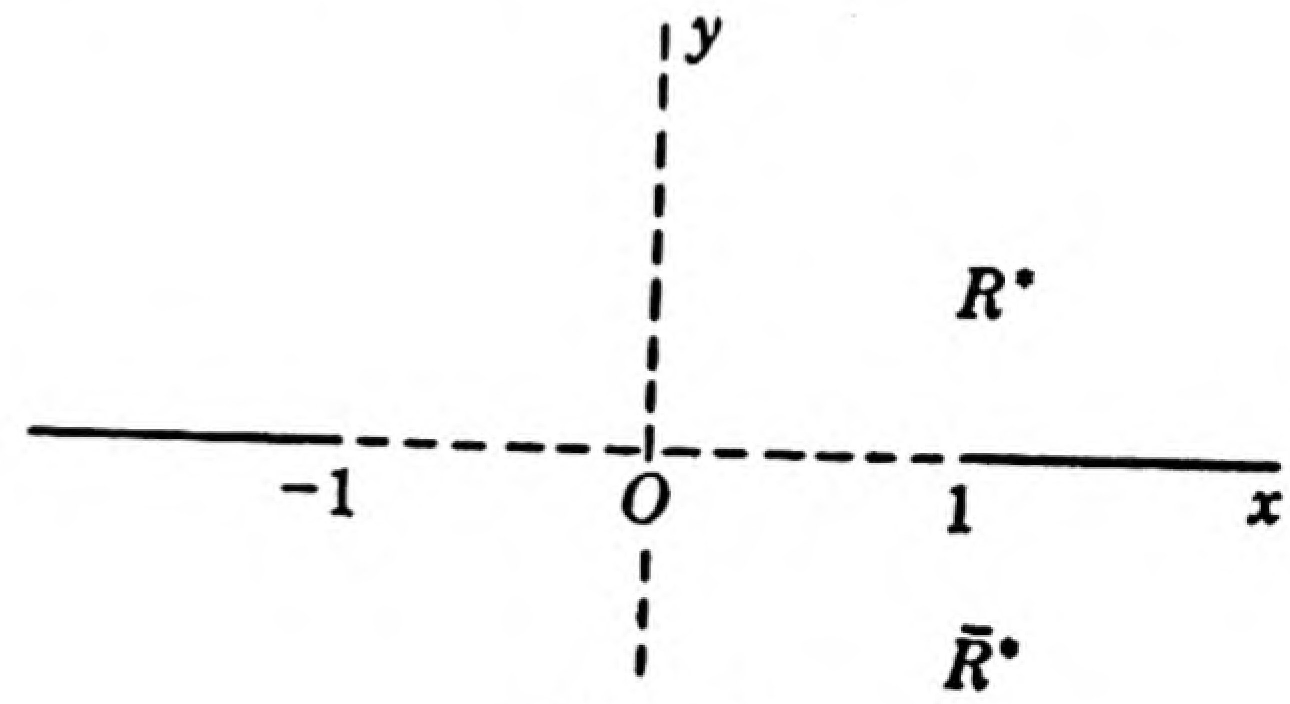


FIG. 9

be the potential energy under a unit moment density across the gap. Then simple dimensional considerations show that

$$\text{(24)} \qquad \begin{aligned} \delta V_n^{*+} &= \epsilon^4 f_n^2 v^+, \\ \delta V_n^{*-} &= \epsilon^2 m_n^2 v^-. \end{aligned}$$

Our formulas (A) and (B) now state that, correct to the lowest order in ϵ, the changes in the squares of the angular frequency are

$$\text{(25)} \qquad \begin{aligned} \delta\lambda_n^+ &= \lambda_n' - \lambda_n^+ = \epsilon^4 \left[\frac{\partial^3 q_n'}{\partial y^3}\right]_0^2 v^+, \\ \delta\lambda_n^- &= \lambda_n' - \lambda_n^- = \epsilon^2 \left[\frac{\partial^2 q_n'}{\partial y^2}\right]_0^2 v^-. \end{aligned}$$

Subtracting these relations, we have

$$\text{(26)} \qquad \lambda_n^+ - \lambda_n^- = \epsilon^2 \left[\frac{\partial^2 q_n'}{\partial y^2}\right]_0^2 v^- - \epsilon^4 \left[\frac{\partial^3 q_n'}{\partial y^3}\right]_0^2 v^+.$$

Apart from the exceptional case when the second normal derivative $\partial^2 q_n'/\partial y^2$ vanishes at the point of the boundary where the gap is located, the second term in (26) is negligible compared with the first for small values of ϵ. Thus λ_n^- is lower than λ_n^+. Physically our result shows that, for small gap lengths, the coupling between the regions R and $\bar{R}$ is due essentially to the rigidity of the plate, the gap acting as a fulcrum for a lever action. Pushing the region R down tends to cause the region $\bar{R}$ to rise.

6. The membrane and acoustic vibrations. The membrane problem of Figs. 1 and 2 is similar to and simpler than the plate problem. For the membrane, $\lambda^- = \lambda'$ rigorously, and

$$\text{(27)} \qquad \delta\lambda_n^+ = \lambda_n' - \lambda_n^+ = \epsilon^2 \left[\frac{\partial q_n'}{\partial y}\right]_0^2 v.$$

In this case it is easy to find an explicit expression for the constant v, since the static Green's function for the region of Fig. 9 can be obtained by conformal mapping.

Our formula (A) has an interesting acoustical application. Consider two rooms separated by a wall. If a small hole is cut in the wall, the frequency

changes can be computed. In this case one of the new frequencies is much lower than all of the original frequencies. This vibration can be regarded as a perturbation of a trivial zero-frequency mode of the two uncoupled rooms.

Bibliography

1. R. J. Duffin, *Nodal lines of a vibrating plate*, J. Math. Phys. vol. 31 (1953) pp. 294–299.

2. R. J. Duffin and D. H. Shaffer, *On the vibration of a ring-shaped plate*, Carnegie Inst. Technol. Tech. Rep. 16, prepared under Contract No. AF-33(616)-1, Department of the Air Force (July, 1952).

3. G. Szegö, *On membranes and plates*, Proc. Nat. Acad. Sci. U.S.A. vol. 36 (1950) pp. 210–216.

Carnegie Institute of Technology,
Pittsburgh, Pa.

INDEX

J

K

L

M

N

O

P